submitted to *Geophys. J. Int.*

Optimal resolution tomography with error tracking and the structure of the crust and upper mantle beneath Ireland and Britain

Raffaele Bonadio[1⋆], Sergei Lebedev[1], Thomas Meier[2], Pierre Arroucau[3], Andrew J. Schaeffer[4], Andrea Licciardi[5], Matthew R. Agius[6], Clare Horan[1], Louise Collins[1], Brian M. O'Reilly[1], Peter W. Readman[1], and the Ireland Array Working Group[†]

[1] *Dublin Institute for Advanced Studies, Dublin, Ireland*

[2] *Christian-Albrechts-Universität zu Kiel, Inst. for Geosciences, Kiel, Germany*

[3] *EDF-DIPNN-DI-TEGG, Aix-en-Provence, France*

[4] *Geological Survey of Canada, Sidney, Canada*

[5] *Université Côte d'Azur, Geoazur, Sophia Antipolis, France*

[6] *Dipartimento di Scienze, Università Roma Tre, Roma, Italy*

Received Year Month Day; in original form Year Month Day

SUMMARY

The classical Backus-Gilbert method seeks localized Earth-structure averages at the shortest length scales possible, given a dataset, data errors, and a threshold for acceptable model errors. The resolving length at a point is the width of the local averaging kernel, and the optimal averaging kernel is the narrowest one such that the model error is below a specified level. This approach is well suited for seismic tomography, which maps three-dimensional Earth structure using large sets of seismic measurements. The continual measurement-error decreases and data-redundancy increases have reduced the impact of random errors on tomographic models.

Systematic errors, however, are resistant to data redundancy and their effect on the model is difficult to predict. Here, we develop a method for finding the optimal resolving length at every point, implementing it for surface-wave tomography. As in the Backus-Gilbert method, every solution at a point results from an entire-system inversion, and the model error is reduced by increasing the model-parameter averaging. The key advantage of our method stems from its direct, empirical evaluation of the posterior model error at a point. We first measure inter-station phase velocities at simultaneously recording station pairs and compute phase-velocity maps at densely, logarithmically spaced periods. Numerous versions of the maps with varying smoothness are then computed, ranging from very rough to very smooth. Phase-velocity curves extracted from the maps at every point can be inverted for shear-velocity ($V_{\rm S}$) profiles. As we show, errors in these phase-velocity curves increase nearly monotonically with the map roughness. We evaluate the error by isolating the roughness of the phase-velocity curve that cannot be explained by any Earth structure and determine the optimal resolving length at a point such that the error of the local phase-velocity curve is below a threshold. A 3D $V_{\rm S}$ model is then computed by the inversion of the composite phase-velocity maps with an optimal resolution at every point. The estimated optimal resolution shows smooth lateral variations, confirming the robustness of the procedure. Importantly, the optimal resolving length does not scale with the density of the data coverage: some of the best-sampled locations display relatively low lateral resolution, probably due to systematic errors in the data. We apply the method to image the lithosphere and underlying mantle beneath Ireland and Britain. Our very large dataset was created using new data from Ireland Array, the Irish National Seismic Network, the UK Seismograph Network, and other deployments. A total of 11238 inter-station dispersion curves, spanning a very broad total period range (4–500 s), yield unprecedented data coverage of the area and provide fine regional resolution from the crust to the deep asthenosphere. The lateral resolution of the 3D model is computed explicitly and varies from 39 km in central Ireland to over 800 km at the edges of the area, where the data coverage declines. Our tomography reveals pronounced, previously unknown variations in the lithospheric thickness beneath Ireland and Britain, with implications for their Caledonian assembly and for the mechanisms of the British Tertiary Igneous Province magmatism.

Key words: Tomography – Seismic tomography – Computational seismology

1 INTRODUCTION

Seismic tomography uses measurements made on seismograms to produce three-dimensional (3D) models of Earth interior, at scales from local to regional to global (e.g., Aki & Lee 1976; Dziewonski et al. 1977; Nolet 2008). The 3D models are solutions of one or a series of inverse problems. Normally, we would like the models to have the highest possible spatial resolution.

1.1 Resolution

Resolution is a fundamental concept in seismic tomography and other imaging fields. It is generally understood as a measure of the ability of an instrument or an experiment to distinguish adjacent features from one another (e.g., Abbe 1873; Helmholtz 1874; Rayleigh 1896; Feynman et al. 1963; Köhler 1981; Sheriff & Geldart 1995). The specific definitions and the limits of the resolution vary with the data type and the design of the image-forming system, from microscopy and telescopy (e.g., Abbe 1873; Helmholtz 1874; Rayleigh 1896) to spectroscopy (McNaught & Wilkinson 2014) and to reflection seismology (Sheriff & Geldart 1995). The term resolution is used widely and in many ways, and its specific meaning can be ambiguous even within the same field (e.g., den Dekker & van den Bos 1997; Demmerle et al. 2015). We thus start with a brief summary on the usage and with definitions.

The classic work on the subject focussed on the resolving power of a telescope or a microscope for a self-luminous double point (Abbe 1873; Helmholtz 1874; Rayleigh 1896). A point source, such as a star, is broadened by diffraction into a finite-width circle on the image. The Rayleigh criterion states that the minimum separation between two light sources required for them to be resolved as distinct objects is proportional to the wavelength of the wave and inversely proportional to the device's aperture (Rayleigh 1896). The specific limit of resolution predicted by the Rayleigh criterion may not necessarily be reached in practice due to noise, depending on the experimental conditions (Ronchi 1961).

The smallest resolvable interval is the angular spread in telescopy, the wavelength or wavenum-

⋆ bonadio@cp.dias.ie

† https://www.dias.ie/ireland_array

ber difference in optical spectroscopy and the spatial distance in seismic imaging. In all cases, an actual point source or point anomaly broaden into a finite-width feature on the image (Fig. 1), so that similar natural definitions of the resolution apply (e.g., Feynman et al. 1963). In reflection seismology, for example, the quarter-wavelength resolution criterion is the equivalent of and is often referred to as the Rayleigh criterion (Sheriff & Geldart 1995).

Seismic tomography differs from telescopy or microscopy in that the image is computed, rather than observed. The resolution, understood as our ability to distinguish features on the image, generally has a complex, non-linear dependence on the data sampling and errors in the data. In their pioneering early work, Backus & Gilbert (1968, 1970) developed a method for computing accurate localized averages of the Earth structure at the shortest length scales that a given dataset, with its given errors, can resolve at given points. Their method determines, for a point $\mathbf{r}_0$, an optimal averaging kernel that is most nearly like $\delta(\mathbf{r}_0 - \mathbf{r})$ (Backus & Gilbert 1970).

Backus & Gilbert (1968) defined the resolving length as the width of the peak of the optimal averaging kernel. In their treatment of errors, Backus & Gilbert (1970) assumed that the variance matrix of the measurement errors can be estimated, and computed the statistics of the resulting model errors using the statistics of the data errors and error-propagation theory.

In this paper, we define the resolving length as the full width at the half-maximum of an averaging kernel at a point (Fig. 1). The averaging kernel is estimated using a point-spread function at the point, computed in a test inversion with the only anomaly being a spike anomaly at this point and with the inversion formulated exactly as that of the real data (Yanovskaya 2005; Oldenborger & Routh 2009; Fitchner & van Leeuwen 2015; Celli et al. 2020). These empirical averaging kernels describe the spreading of spike anomalies $\delta(\mathbf{r}_0 - \mathbf{r})$ in space. The shape of the kernels is not postulated a priori. With our inversion set-up, the kernels turn out to have a natural bell shape, close to a Gaussian in cross-section (Section 3.3). This resolving length definition is similar to that using the half-width of a cone, illustrated and applied in a number of previous tomography studies (e.g., Barmin et al. 2001; Ritzwoller et al. 2002; Lebedev et al. 2003; Celli et al. 2020).

The resolving length is equal to the distance above which two spike anomalies can be distinguished (Fig. 1). The resolution of the imaging is said to be higher if the resolving length is

shorter, and lower if the resolving length is greater. The term resolution is also often used with the meaning of the resolving (averaging) length—as in, for example, "resolution of 100 km" (Nolet 2008, p. 221).

In linear inverse theory, resolution is often discussed in terms of the closeness of the resolution matrix to the identity matrix. For a linear inverse problem

$$\mathbf{Am} = \mathbf{d}, \tag{1}$$

where $\mathbf{m}$ is the model vector, $\mathbf{d}$ the data vector and $\mathbf{A}$ the sensitivity matrix, the solution $\mathbf{m}$ can be written as

$$\mathbf{m} = \mathbf{A^{-1}d}, \tag{2}$$

where $\mathbf{A}^{-1}$ is a generalized inverse (e.g., Nolet 2008; Menke 2012). Substituting Eq. (1) into Eq. (2),

$$\mathbf{m} = \mathbf{A^{-1}Am^{(true)}} = \mathbf{Rm^{(true)}}, \tag{3}$$

where $\mathbf{R} = \mathbf{A^{-1}A}$ is the model resolution matrix. It can be thought of as a blurring filter through which we see the real Earth ($\mathbf{m^{(true)}}$) on the tomographic image.

If the observed data $\mathbf{d}$ is the sum of the error-free data $\mathbf{d^{(true)}}$ and data errors $\mathbf{e}$, then the error of the solution $\mathbf{m}$ is (Nolet 2008):

$$\mathbf{m - m^{(true)}} = \mathbf{A^{-1}d - m^{(true)}} = \mathbf{(A^{-1}A - I)m^{(true)} + A^{-1}e} = \mathbf{(R - I)m^{(true)} + A^{-1}e}. \tag{4}$$

Two components of the model error are the blurring of the true structure by the tomographic "filter," quantified by the difference of the resolution matrix and the identity matrix $(\mathbf{R} - \mathbf{I})$, and the propagated data errors $\mathbf{A^{-1}e}$.

An additional, third component of the model error comes from the error of the linear relationship $\mathbf{Am} = \mathbf{d}$ itself. If the parameters of the model $\mathbf{m}$ sample a 3D volume or a 2D plane, then the i-th row of the matrix $\mathbf{A}$ defines the sensitivity volume or sensitivity area assigned to the measurement d_i, specifying the weights of the model parameters within this volume or area. The sensitivity volumes depend on the measurement method and on the imperfectly known structure

of the Earth (Dahlen & Tromp 1998) and are, at best, a good approximation. Errors may also result from the problem linearization itself (e.g., Rawlinson et al. 2010).

If the off-diagonal elements of $\mathbf{R}$ are all zero, then each parameter is determined uniquely. In practice, they are usually non-zero, so that the parameters are weighted averages of the true model parameters. One measure of the resolution $\mathbf{R}$ is the net size, or the spread, of the off-diagonal elements (Menke 2012).

The broadening of a point anomaly in the model (Fig. 1) is described by the leakage from the diagonal to off-diagonal elements of the resolution matrix. The averaging kernel of a tomographic inversion, estimated using a point-spread function, yields an approximation of a row or a column of the model resolution matrix (Ritzwoller et al. 2002; Menke 2015). Point-spread functions computed for every parameter of the model yield an estimate of the entire resolution matrix and a map of the resolving lengths (Ritzwoller et al. 2002; Celli et al. 2020). The definition of resolution in terms of $(\mathbf{R}–\mathbf{I})$ is thus similar to its definition in terms of our ability to distinguish adjacent features. An important difference is that $\mathbf{R}$ is independent of errors in the data and approximations (Menke 2012), whereas the optimal averaging kernels—and our ability to distinguish adjacent features from one another—do depend on the errors (Backus & Gilbert 1970).

1.2 Errors

As the resolving length decreases, the error of the localized average increases (Backus & Gilbert 1970). The trade-off of the resolution and variance is a general principle of inverse theory: the resolution spread can be decreased at the expense of increasing the variance, and vice versa (Menke 2012).

The model variance arises from the incompleteness of the data sampling and from the errors in the data and the methods' approximations. The growth in the coverage of the Earth with seismic stations over the last few decades has driven a continuous increase in the resolution of global and regional tomographic models (e.g., Rawlinson et al. 2010). Errors of most types have also decreased, and the increasing data redundancy is reducing the impact of random, uncorrelated

errors. In many cases, however, the remaining errors are systematic, so that their effect on the models is resistant to the data-redundancy increase.

For example, in tomography using teleseismic delay times or phase delays accumulated between sources and stations, the major source of errors is the uncertainty in the event location and origin time and, for waveform analysis techniques, of the source mechanisms. The source parameters in published catalogues were computed using approximate, assumed Earth structure, and their errors in a given source region tend to be systematic (Lebedev et al. 1997; Bijwaard et al. 1998). Differential measurements, including inter-station, surface-wave measurements as in this study, isolate the information on local structure from the effect of the source but tend to have complex sensitivity volumes, shaped by structural heterogeneity both near and away from the stations (e.g., de Vos et al. 2013) and impossible to map exactly. Measurement errors due to instrumental errors—timing errors, response-correction errors, polarity reversals—remain a problem (e.g., Weidle et al. 2013; De Laat et al. 2019) and are also systematic.

The errors in the data propagate into errors in the models and limit the resolution of the imaging via the resolution-variance trade-off (Backus & Gilbert 1970; Menke 2012). In order to keep the model error below a certain level, the resolving length must be sufficiently large. Comparisons of different global tomographic models give a vivid illustration of this trade-off: the models show an excellent mutual agreement at long spatial wavelengths but a progressively decreasing agreement at decreasing wavelengths (e.g., Boschi & Dziewonski 1999; Becker & Boschi 2002; Schaeffer & Lebedev 2015; Schaeffer et al. 2016).

1.3 Spatially variable resolution

The optimal averaging-kernel width (Backus & Gilbert 1968) varies spatially because of the unevenness of the data sampling and error distribution. The unevenness of the data coverage is due to the irregular geographical distribution of the sources and receivers.

A number of approaches have been implemented to accommodate the spatially variable level of detail in model (Rawlinson et al. 2010). Irregular parameterizations aim to place the nodes of the grid only where they are required by the data, or have the size of the blocks scaled with data

sampling (e.g., Chou & Booker 1979; Tarantola & Nercessian 1984; Fukao et al. 1992; Sambridge et al. 1995; van der Hilst et al. 1997; Bijwaard et al. 1998; Bijwaard & Spakman 2000; Debayle & Sambridge 2004; Zhao 2004; Sambridge & Rawlinson 2005). The grid density can vary according to chosen indicators of the resolving power of the data at different locations, such as hit counts or sensitivity-matrix column sums. The inversion can be parameterized using nested grids, with a higher-resolution grid where the sampling is greater or in the area of primary interest—for example, a denser grid for a region embedded into a sparser global grid. Adaptive parameterization aims to adjust in the course of the inversion, matching the spatially varying structural information yielded by the data (e.g., Michelini 1995; Curtis & Snieder 1997; Sambridge & Faletič 2003; Rawlinson & Kennett 2004). Recently, dynamic parameterizations in Bayesian frameworks without explicit regularization have been developed and applied to tomographic and other seismic-data inversions (e.g., Bodin & Sambridge 2009; Piana Agostinetti & Malinverno 2010; Bodin et al. 2012; Piana Agostinetti et al. 2015; Galetti et al. 2016; Hawkins et al. 2019).

Commonly used forms of regularization of tomographic inversions with local parameterisations are norm damping and smoothing, with the smoothing implemented, typically, as gradient damping or Laplacian damping or both (e.g., Nolet 2008; Lebedev & van der Hilst 2008). Lateral and radial smoothing increases the width of the averaging kernels and decreases the nominal resolution of the models. The resolution varies spatially even in inversions with constant factors of regularization, because the same regularization term has a smaller effect where the data coverage is denser and the data-misfit term is greater (e.g., Nolet 2008). The spatially varying resolution can be examined using resolution tests (e.g., Rawlinson & Spakman 2016) or resolution-matrix calculations (e.g., Boschi 2003; Deschamps et al. 2008), but neither of the approaches can quantify the impact of unknown, correlated errors in the data.

1.4 Optimal resolution tomography

Backus & Gilbert (1970) formulated and solved the problem of how to find the shortest length scale over which local average structure at a particular point can be determined with the variance

under a specified amount. In other words, they determined the optimal resolving length (the width of the peak of their optimal averaging kernel), or optimal resolution, given the errors.

The Backus-Gilbert averaging kernels were used in a number of mantle tomography studies (e.g., Trampert & van Heijst 2002). Beyond that, their work has been a major influence on the geophysical inverse theory (e.g., Chou & Booker 1979; Tarantola & Nercessian 1984; Parker 1994; Nolet 2008; Menke 2012) and inspired the development of optimally localized average (OLA) methods in other fields. The computationally efficient Subtractive Optimally Localized Averages (SOLA) method (Pijpers & Thompson 1994), popular in helioseismology, has recently been adapted to seismic tomography by Zaroli (2016, 2019) and Zaroli et al. (2017).

In this study, we pose the problem in the same way as Backus & Gilbert (1968, 1970): how can we find the solution of the tomographic inverse problem with the smallest resolving length at every point, such that the error of the local average at the point is below a specified threshold? We recognise that the statistics of the errors in the data and approximations are unknown. From what we know about the errors, we can say that they are unlikely to have a zero mean and are substantially correlated, in ways that are unlikely to be guessed or modelled accurately. This makes them difficult to handle using error-propagation estimation used in the linear inverse theory (Nolet 1985).

But what if we had a way to evaluate the posterior model error at a point directly? With that, we could solve the inverse problem repeatedly, adjusting the width of the averaging kernel at this point until the error is just below the specified threshold. In other words, we would be able to determine the optimal resolution at the point. Similarly to the Backus-Gilbert method, this would reduce the variance by increasing the volume over which the model parameter is averaged, until the error is acceptable. Like in the Backus-Gilbert method, every estimate at a point would require a full inversion of the entire system—a series of inversions, in fact.

The key advantage of this approach over the existing ones stems from its key ingredient, the direct evaluation of the model error. In the following, we develop the optimal resolution tomography for the surface-wave tomography problem, set up as a sequence of phase-velocity tomography and the point-by-point inversion of local phase-velocity curves. We shall start with presenting our

large regional dataset, introduce the phase-velocity measurements and, then, describe the implementation and validation of the optimal resolution tomography, with an application to the imaging of the crust and upper mantle beneath Ireland and Britain.

1.5 Imaging Ireland and Britain

The lithospheric evolution of the Ireland-Britain region (Tiley et al. 2004; Landes et al. 2007; Holland & Sanders 2009; Davis et al. 2012; Cogné et al. 2016) and the mechanism of its enigmatic Paleogene intraplate volcanism (White & Lovell 1997; Jones et al. 2002; Al-Kindi et al. 2003) are poorly understood, in large part due to the lack of information on the region's lithospheric structure. Much of our present knowledge of the seismic structure of the crust beneath Ireland and Britain is from active source seismic refraction and reflection experiments (Bamford et al. 1978; Edwards & Blundell 1984; Bott et al. 1985; Jacob et al. 1985; Freeman et al. 1988; Lowe & Jacob 1989; Snyder & Flack 1990; Klemperer & Hobbs 1991; Klemperer et al. 1991; Barton 1992; O'Reilly et al. 1996; Masson et al. 1998; Landes et al. 2000; Hodgson 2001; Kelly et al. 2007; O'Reilly et al. 2010; Maguire et al. 2011; O'Reilly et al. 2012). Maps of the Moho depth have been obtained from the results of the active-source experiments and, also, by combining them with those from receiver functions (e.g., Chadwick & Pharaoh 1998; Asencio et al. 2003; Landes et al. 2005; Kelly et al. 2007; Di Leo et al. 2009; Davis et al. 2012; Licciardi et al. 2014, 2020).

Passive-source investigations of the region's crust and upper mantle included teleseismic travel-time comparisons (Masson et al. 1999), teleseismic body-wave tomography (Arrowsmith et al. 2005), local earthquake tomography (Hardwick 2008), receiver-function (Shaw Champion et al. 2006; Tomlinson et al. 2006; Landes et al. 2007; Davis et al. 2012; Licciardi et al. 2014) and shear-wave splitting (Do et al. 2006; Bastow et al. 2007) analysis and surface-wave tomography (Polat et al. 2012; Nicolson et al. 2012, 2014; Galetti et al. 2016). Magnetotelluric and gravity studies provided additional, complementary information (Brown & Whelan 1955; Readman et al. 1997; Rao et al. 2007). Recently, petrological modelling and inversion were applied to integrate seismic, geothermal, compositional, and magnetic data (Fullea et al. 2014; Jones et al. 2013; Mather et al. 2018; Baykiev et al. 2018; Mather & Fullea 2019).

The coverage of Ireland with seismic stations was sparse and uneven until recently. Regional surface-wave studies to date (Polat et al. 2012; Nicolson et al. 2012, 2014; Galetti et al. 2016) focussed on parts of the region and used data in limited period ranges. Continent-scale tomographic models typically include Ireland and Britain at the edge of the model, imaged with relatively low resolution (e.g., Marquering & Snieder 1996; Fry et al. 2008; Schivardi & Morelli 2009; Rickers et al. 2013; Soomro et al. 2016).

2 DATA AND MEASUREMENTS

In this study, we used the abundant, newly available data in the Ireland-Britain region in order to obtain numerous phase-velocity measurements in very broad frequency ranges (from periods as short as 4 s to those as long as 500 s) and to image the entire region at a new level of detail. We used phase-velocity measurements, generally more accurate than group-velocity ones (e.g., Meier et al. 2004; Boschi et al. 2013; Soomro et al. 2016) and yielding more accurate maps (Dahlen & Zhou 2006). The phase-velocity dispersion curves were obtained for 11238 two-station paths across the area using a combination of a recent implementation of the two-station method and waveform inversion. Our dataset includes all the data recorded by the broadband networks in Ireland and all the publicly available data from the broadband stations in Britain (Appendix A1). In total, our measurements were made on data recorded between 1981 and 2018, but the bulk of the data is from the last decade. Thanks to the recent growth in the number of stations, especially in Ireland, our dataset provides an unprecedentedly dense data coverage of the entire region (Figs. 2, 3).

The waveform data went through automated quality checks and preprocessing. The integrity of the data was ensured by removing all the data that were incomplete, clipped, had gaps. The seismograms were converted to displacement by the removal of the instrument response and down-sampled to 1 Hz.

2.1 Teleseismic two-station cross-correlation

For each available station pair, we searched the Global Centroid Moment Tensor catalogue (Dziewonski et al. 1981; Ekström et al. 2012) for teleseismic events within the operating time period of the

two stations and with a chosen back-azimuth range of $\pm 5^{\circ}$ from the station-station great circle path (GCP). Events with a moment magnitude greater than 4.9 were chosen, using a distance-dependent magnitude threshold (Schaeffer & Lebedev 2013). Only Rayleigh-wave data from the vertical component were used for the measurements; Love wave measurements will be incorporated in a future study.

The use of the two-station method, as introduced by Sato (1955), in surface-wave analysis allows us to compute phase-velocity dispersion of the surface waves that travel approximately along the GCP between stations of a pair. It is possible to make a regional investigation using teleseismic earthquakes, since the phase effects of the source (earthquake) and the common path between the source and the receivers cancel out (e.g., Meier et al. 2004; Soomro et al. 2016). In practice, the waves from a given earthquake do not travel exactly along the GCP, so we need to allow a certain tolerance for the alignment between the earthquake and the pair of stations. This tolerance can be chosen depending on the availability and quality of the data. In this paper, the low threshold of $\pm 5^{\circ}$ is chosen, due to high availability of recordings and relatively high signal-to-noise ratios. Such conservative data-selection approach is allowed by the enormous quantity of available waveforms, in contrast to other studies in which the threshold has to be higher, due to a smaller quantity of available data and a lower signal-to-noise ration of recordings (e.g., Bonadio et al. 2018).

Although the tolerance around the station-station GCP is very small, one may argue that the misalignment may have a non-negligible effect on the accuracy of the calculated phase velocity. However, the imperfect alignment of the two stations and the event has no immediate effect on the measurement accuracy because the phase velocities are computed from the phase of the cross-correlation function and the difference between the distances from the event to each of the stations, rather than the interstation distance between the two-stations (e.g., Soomro et al. 2016).

In this work we use the implementation of the two-station method by Meier et al. (2004). The automated measurement procedure is adapted for our particular dataset from Soomro et al. (2016). For each teleseismic event, the vertical component seismograms recorded at the two stations are cross-correlated. The cross-correlation signal is then filtered using a frequency-dependent Gaus-

sian band-pass filter, so as to minimize the effect of noise and interferences on the fundamental mode. The resulting signal is then weighted in the time domain to reduce the effects of scattering and higher modes. The phase velocity is computed from the resulting signal in the Fourier domain as the arctangent of the ratio of the imaginary to real part of the Fourier spectrum. This approach works best if the fundamental mode is dominant compared to any other type of signal or noise in its vicinity on the traces.

The accuracy of the resulting curve depends on the amplitude of the fundamental mode content, the signal-to-noise ratio, unmodelled surface-wave diffraction, and the interferences of the Rayleigh and Love, fundamental and higher modes. The smaller the errors due to the diffraction and interferences, the more accurate and the smoother the dispersion curve. To minimize the effect of the errors in the curves on the final, average measurements, we only accept smooth portions of phase-velocity curves. We also exclude the outlier measurements and, also, accept only the curves not unrealistically far from a pre-calculated reference dispersion curve (how the reference curves used in this work are computed is discussed in Section 2.3). The accepted phase-velocity segments are selected subject to the following criteria (e.g., Soomro et al. 2016):

- segments that present a low number of samples are not selected for the final averaging; due to the logarithmic sampling of the frequency axis the minimum acceptable length of the segments varies with period (i.e., longer segments are accepted at longer periods, while shorter segments may be accepted at shorter periods);
- for each pair of stations, the ensemble of selected curves is analysed for the number of measurements at each period; if a minimum number of measurements (10 in this study) has not been reached, the measurements at those periods are not included into the final dataset;
- if a systematic inconsistency between measurements from events at opposite directions from the station pair is detected, the measurements from the station-pair are removed from the dataset. The procedure computes the average phase velocity and the standard deviation for each set of measurements corresponding to the two directions, and the measurements are rejected if the phase-velocity difference for the two directions exceeds a certain threshold. It is important not to introduce this inconsistency in the data, as this effect could indicate instrumental errors (station timing

or instrument response) or strong diffraction effects. In the dataset in this study, the procedure has not detected any station pair displaying such inconsistency.

We used all station pairs with interstation distances greater than 1 km and within the area of the map in Figs. 2 and 3. For each station pair, we computed an average over, typically, many tens to a few hundreds of one-event measurements, made using recordings of earthquakes in different source regions, in different directions from the station pair. The averaging over a large number of measurements reduces the effect of errors due to diffraction and interferences between the fundamental and higher modes and results in robust average measurements in very broad period ranges. The automatic selection described applies to the measurements obtained with both the two-station cross-correlation (CC) and the automated multimode inversion (AMI) (Section 2.2).

2.2 Measurements from waveform inversion

We use the Automated Multimode Inversion of surface and S-wave forms (Lebedev et al. 2005) to complement our phase-velocity measurements at intermediate and long periods. The AMI method simultaneously fits S, multiple S and surface waves for each source-station pair, using synthetic seismograms generated by mode summation in seismogram-dependent time-frequency windows. As a by-product of the waveform inversion, it measures phase velocities within the period bands constrained by this particular waveform fit. For each source-receiver pair, we extracted the fundamental-mode phase velocities. We then used pairs of stations at the same azimuth from the event ($\pm 5^\circ$) to calculate the interstation phase velocity using the relation

$$c_{12} = \frac{\Delta_2 - \Delta_1}{\Delta_2/c_2 - \Delta_1/c_1}, \tag{5}$$

where $c_{i=1,2}$ and $\Delta_{i=1,2}$ are the phase velocity and the distance between the source and each of the two stations, respectively, and c_{12} is the interstation phase velocity (Lebedev et al. 2006; Agius & Lebedev 2014). The advantage of using AMI for interstation measurements and combining them with those obtained by cross-correlation is that the waveform inversion can measure phase velocities at long periods, where the cross-correlation often fails because of the strong interference

between surface and body waves (e.g., Meier et al. 2004; Zhang et al. 2009; Agius & Lebedev 2013).

With these two automated methods, we computed a very large number of phase-velocity, inter-station curves in a total period range of 4–500 s (Fig. 4). The measurements generated by the different techniques (at short, intermediate and long periods—by cross-correlation, and at intermediate and long periods—by waveform inversion) are consistent where they overlap. The curves from cross-correlation and AMI are averaged all together (Section 2.1). The final set of measurements used for the construction of the velocity maps (Section 3) is shown in Fig. 4. The period range of the final dataset is 4–500 s.

2.3 Reference model for the measurements

A $\pm 2\pi$ ambiguity arises when phase velocities are computed from the cross-correlation function (e.g., Meier et al. 2004; Soomro et al. 2016; Bonadio et al. 2018; El-Sharkawy et al. 2020). We need a reference model to discriminate the curves that represent the true Earth structure from the ones that are shifted up or down by the trigonometric ambiguity. For longer periods, identifying the correct dispersion curve is normally straightforward (Fig. 5 d). The ambiguity may occur, instead, at higher frequencies, where the choice of the correct curve is not always trivial, with the curves close to each other. Using an accurate a priori reference phase-velocity curve is thus important for the phase-velocity measurements. It would not work, for example, to use a global reference model such as AK135 (Kennett et al. 1995) or PREM (Dziewonski & Anderson 1981) if the study area presents substantially different phase-velocity dispersion compared to these models.

We computed a reference model for the region using the method of Bonadio et al. (2018), which provides a data-based initial reference curve for the area. Stacking together all possible phase-velocity curves derived from cross-correlations for the entire set of station pairs, without any selection applied, we produce a density plot, as in the bottom panel in Fig. 6. Applying a loose selection on the data (such that for each measurement we only plot the curve closest to AK135 model in a certain period range, 20–50 s in this study) we can improve the density plot (top panel in Fig. 6) and obtain an average phase-velocity curve for the region that can then be used as the

reference for the final, more precise, one-by-one, phase-velocity measurements. The reference curve is obtained from the maximum values of the density distribution and then smoothed by means of a very weakly regularized inversion for a shear-velocity profile (Section 4.2).

Provided that a large number of measurements are contributing to the stack, one could analyse parts of the data for different sub-areas, obtaining different reference curves for them, in case the study area presents strong heterogeneity. In this study, we used five sub-areas, with average phase-velocity curves within them relatively similar to each other (Fig. 7). The average obtained from this stacking procedure is a useful reference for the measurements that follow but should not be taken as representing the real Earth structure, because systematic errors due to diffraction, scattered waves, the interference between the fundamental and higher modes and noise may bias the stack. Errors due to these effects are reduced by our strict measurement selection, described in Section 2.1. Although we obtained different reference curves for the five different sub-regions, the small differences between them did not justify using different reference curves in this study. Generally, one would use different reference curves if the study area comprises more diverse tectonic settings.

3 OPTIMAL RESOLUTION PHASE-VELOCITY MAPS

We wish to build a set of phase-velocity maps, at many periods within the range of the measurements, with an optimal resolving length at every point. The optimal resolving length is the smallest width of the averaging kernel such that the model error at the point is below a specified level (Backus & Gilbert 1968, 1970). The model error depends both on the data sampling and on the errors in the data and in the methods' approximations. According to the resolution-variance trade-off, the model error is expected to increase with the decreasing width of the averaging kernel (Backus & Gilbert 1970; Menke 2012).

The width of the averaging kernel at a point in our maps depends on the smoothing applied to the inversion: inversions with stronger smoothing produce models with broader averaging kernels. The optimal averaging kernel from the range given by the different levels of smoothing can be selected if we can evaluate the model error at the point. Here we propose a method for estimating

the errors of the sets of phase-velocity maps at every point. This amounts to tracking the errors to this key point of the tomographic scheme and provides a means to determine the optimal resolution at every point such that the effect of the errors is acceptably small.

3.1 Phase-velocity maps with different smoothness

We invert phase-velocity curves from all interstation pairs for phase-velocity maps using a least-squares technique, LSQR (Paige & Saunders 1987), with smoothing, which is the primary means of regularization, and weak norm damping (Lebedev & van der Hilst 2008; Deschamps et al. 2008; Darbyshire & Lebedev 2009; Pawlak et al. 2012). The maps are parameterized using a triangular grid with a 10 km knot spacing. The five model parameters at each grid knot include the isotropic-average anomaly and four anisotropic coefficients, two for π-periodic and two for $\pi/2$-periodic variations with azimuth (e.g., Smith & Dahlen 1973; Deschamps et al. 2008). The inversion at a period solves the system of equations yielded by all the path measurements at this period for the isotropic and anisotropic terms. Regularization is by means of norm damping (which penalizes model-parameter amplitudes), Laplacian smoothing (which penalizes the difference between the anomaly at a node and the average anomaly over this and the nearest neighbouring nodes), and gradient damping (which penalizes the differences between pairs of neighbouring model knots), all of which are applied independently to the isotropic and anisotropic components of the model (Lebedev & van der Hilst 2008; Endrun et al. 2008).

In the final, composite tomographic maps (Section 3) the solution at each of the 4328 grid knots comes from a whole-system inversion generally different from the one for the neighbouring points, with its own level of regularization chosen to yield an optimal local resolution. First, a series of 2D inversions for phase-velocity maps is performed at each period, with the smoothing coefficients incremented at small steps from very low to very high. The gradual change in smoothing (S) for the velocity maps (Fig. 8) is obtained with the Laplacian smoothing coefficient three times the gradient damping coefficient, norm damping as small as possible (chosen empirically so as to be able to suppress localised artefacts near the locations of some of the stations, for example, but also to be small enough not to affect the amplitude of the anomalies elsewhere across the maps) and the

regularization coefficients for the anisotropic terms 1.5 times those for the isotropic term. After an initial inversion, each phase-velocity map is recomputed with 25% of the "outlier" measurements discarded at each frequency (Lebedev & van der Hilst 2008). The outliers are defined here as the measurements fit the worst by the model; the procedure effectively selects the most mutually consistent measurements and removes the least mutually consistent ones, likely to contain the largest errors in the dataset. Although this is effective in reducing noise in the dataset (e.g., Lebedev & van der Hilst 2008; Schaeffer & Lebedev 2013; Celli et al. 2020), the remaining data still have errors, and it is these that errors translate into model errors and, together with the incompleteness of data sampling, impose lower limits on the optimal resolving lengths.

Once the 2D tomographic maps are produced at each period (47 logarithmically spaced periods from 5.1 s to 454.1 s) and for each smoothing level $S_{\mathrm{i}=0,1,2,\ldots,38}$, we extract phase-velocity curves at each of the 4328 knots of the grid (a total of 168,792 dispersion curves). The sample spacing is chosen to vary logarithmically with period so as to roughly balance the information contained within the different parts of the phase-velocity curve, sensitive to different depth intervals within the Earth (e.g., Lebedev et al. 2013). In the top panels of Fig. 9, we plot, as examples, the dispersion curves extracted at two knots of the grid, one in Wales (left) and one in Ireland (right).

3.2 Optimal smoothness at a point

As we show below, a local phase-velocity curve at a point—extracted from a set of phase-velocity maps at different periods—generally has errors that vary in concert with the roughness of the maps: the smoother the maps (the smaller the roughness), the smaller the errors. This offers us a straightforward way of finding the optimal resolving length at a point: the averaging kernel width depends on the smoothing, and it is optimal when the estimated model error is just below a threshold. The problem is now reduced to finding the smallest level of smoothing of the maps such that the errors of the local dispersion curve are below the threshold. The procedure is repeated for each point. The optimal smoothness of the full inversion generally varies from one point to another.

To identify the optimal S_{i} for each knot of the model, we use the following strategy. For ev-

ery local dispersion curve—for each S_{i} and at each knot—we estimate its errors by isolating its roughness. Due to the surface-wave sensitivity kernels' broad depth range and smooth variations with period, any realistic phase-velocity curve is smooth. This is true even for dispersion curves computed for unrealistic Earth models with highly oscillatory depth dependence of seismic velocities. The rough (not smooth) variations of phase velocities with period are, therefore, entirely due to errors. The frequency-dependent roughness of a phase-velocity curve can thus yield an estimate of its frequency-dependent errors (Ravenna et al. 2018).

The roughness of a phase-velocity curve can be isolated by means of a very weakly-regularized inversion of it for a 1D Earth model. The smooth component of the curve can be matched closely by a synthetic curve computed for a best-fitting 1D Earth model (which is not required to be realistic in this inversion and can be oscillatory). The rough component is then given by the remaining misfit—varying rapidly with period—between the curve and its synthetic counterpart. This rough component cannot be fit by any Earth structure and is due to the errors of the dispersion curve. An estimate of the period-dependent error can now be obtained from the misfit or its envelope (Ravenna et al. 2018). This error estimate is conservative in the sense that the weakly regularized inversion fits the dispersion curve as closely as possible, even if this requires an unrealistic, oscillatory 1D model. This may be offset, to some extent, by the fact that this approach will miss errors that do not vary with period or vary with period slowly and, thus, do not manifest themselves in the dispersion-curve roughness. Event mislocations, for example, could cause frequency-independent errors in source-station measurements, but they would not have a significant effect on inter-station measurements. In the inter-station measurements as used here, an instrument-response error at one of the two stations could produce a measurement error with a weak frequency dependence, but this would be likely identified by our routine comparisons of the measurements using sources in the two different directions from the station pair. For any remaining errors in the data to translate into frequency-independent errors in the phase-velocity curves, they would need to cause the same bias in the phase-velocity maps in the same location at different periods. The largest source of remaining errors is probably unmodelled surface-wave diffraction. Fortunately for our purposes here, it is strongly frequency-dependent (e.g., Meier et al. 2004; Kolínskỳ et al. 2021), which is

manifest in the curve roughness. For these reasons, the frequency-independent errors are likely to be small.

Our weakly regularized inversion for a 1D shear velocity profile is a non-linear, Levenberg-Marquardt gradient search (e.g., Meier et al. 2004; Lebedev et al. 2006; Endrun et al. 2008; Erduran et al. 2008; Agius & Lebedev 2013, 2014) (see Section 4 for details on the inversion algorithm). The misfit is computed as the relative data-synthetic misfit integrated along the length of the curves. With all the 1D inversions performed using the same weak regularization, we find that the relative data-synthetic misfit is smaller for curves extracted from maps with higher $S_{\rm i}$, and higher for curves extracted from maps with lower $S_{\rm i}$ (Fig. 9). These inversions are not meaningful in terms of real Earth structure; we use them only to isolate the roughness of the curves, which is entirely due to errors.

Extremely smooth phase-velocity maps are the most robust and accurate, at their spatial wavelengths, but at the cost of lower resolution: they display large-scale structural trends but fail to show structure at a high level of detail. Conversely, phase-velocity maps that are not smooth enough will fit noise and may be dominated by artifacts.

Figs. 10, 11, and A1 show that the error of the phase-velocity curves, estimated by the roughness-isolating 1D inversions, decreases with increasing smoothing of the 2D velocity map. Fig. 10 shows the root mean square (RMS) misfit (panels (e) and (f)) computed for every inversion at two different sets of knots, in Ireland (top panel) and in Britain (bottom panel), as indicated by the black dots in the maps. Panels (a) and (b) show the phase-velocity curves extracted from the 2D tomographic maps with three different level of smoothing (S), as indicated by the colours (red, green, and blue, respectively for "low", "intermediate" and "high" S). The relative misfit is shown in panels (c) and (d). It is clear from this figure that the smoother the 2D phase-velocity map, the smaller the misfit in the 1D inversion for shear velocity. Importantly, the estimated error increases nearly monotonically with the decrease of the smoothing coefficient, which also confirms that the models converge consistently, unaffected by any local minima, for example. The same behaviour is observed in Fig. A1, where the portion of the phase-velocity curves with estimated errors less than 0.15%, rather than the RMS misfit, is plotted as a function of S.

The data sampling given by our large phase-velocity dataset is so redundant that random noise largely cancels out, but not the systematic errors, which may be due to wave propagation effects and instrumental problems. This is apparent from the fact that the error of the local phase-velocity curve does not scale with the data sampling (Fig. 12, Section 3.3).

We now define an empirical criterion to identify the optimal S_i for each knot of the model. We set a threshold of 0.15% for the relative misfit and accept only those phase-velocity curves that produce misfit within this threshold at at least 75% of the periods. This way, curves too rough (low S_i) are discarded, and we can select the roughest of the remaining, smoother curves as the optimal one, with the corresponding optimal level of smoothing and the corresponding optimal averaging kernel. The criteria for choosing the error threshold have been chosen empirically after extensive testing.

In Fig. 13, the composite phase-velocity maps constructed using our preferred threshold for the acceptable errors of the local dispersion curves (c) are compared with two composite maps produced with higher thresholds and two composite maps with lower thresholds—that is, two looser and two stricter selections (columns (a), (b) and (d), (e) respectively)—for three different periods. Our empirical threshold choice is intended to yield an estimate of an optimal resolving length at every knot. In Fig. 11 we show that the cumulative error, as well as the portion of the curve with estimated error less than 0.15%, are changing with S nearly monotonically. Parts of the maps near the region boundary, where the coverage is extremely low, have been removed from the analysis based on sensitivity-matrix column sums—accepting only the knots with the sum's value over a certain threshold.

The map of the optimal resolving length at 78 s period is shown in Fig. 12, together with a map of the chosen values of S_i and a map of the density of the data coverage. The optimal resolving length does not scale with the density of data sampling. This is, in part, due to the complex azimuthal unevenness of the coverage and, in part, due to the errors in the data—given the substantial redundancy of the dataset, probably due to systematic errors primarily. The optimal values of the smoothing coefficient also do not scale with the data sampling and could not be

estimated directly from the column sums of the sensitivity matrix. This necessitates the sequence of numerous inversions included in our approach.

3.3 Estimating spatial resolution

The procedure used to estimate the resolution length at a point is as follows. We simulate a spike perturbation for this one knot of the model grid, with no anomaly elsewhere, and use our sensitivity matrix $\mathbf{A}$ to determine the synthetic data vector $\mathbf{d}$ (Eq. (1)). We then solve the inverse problem with exactly the same regularisation as applied to the inversion of the data (Fig. 14) and evaluate how much the delta-like perturbation has spread after the inversion. To do so, we register the values of the output anomaly along great circles that extend 4 degrees from the point in each direction and are oriented at densely spaced azimuths covering the full azimuth range (Fig. 15). Smooth curves through the points, determined using cubic splines, display a natural bell shape (Fig. 15, left), similar to a Gaussian function. The half-width of the averaging kernel is computed as an average over the half-widths at the half-maxima (HWHM) of the curves at all azimuths. The resolving length is twice that, that is, the full width at the half maximum.

The procedure is repeated for every knot of the grid and for every smoothing level. The method for resolving-length estimation can break at the edge of the region, where the data sampling deteriorates. Where the HWHM cannot be defined (e.g., the amplitude of the anomaly is too low, or its width too high), the resolution length is set to infinite. Maps of the estimated resolving length at different periods obtained with this method are shown in Fig. 16.

The optimal-resolution phase-velocity maps, for a selection of periods, are shown in Figs. 17 and 18. It is not possible to identify the errors in the measurements going into the tomographic inversion, but we can evaluate them between the phase-velocity map step and the 1D inversion step, at which point the optimal resolution is determined. The estimated optimal resolution shows smooth lateral variations, confirming the robustness of the procedure (Fig. 16).

Fig. 19 illustrates the advantages of using the optimal resolution scheme (1^{st} column) compared to inversions with a constant smoothing factor (2^{nd}, 3^{rd}, and 4^{th} columns). Using a uniform (rough, medium, or smooth) regularization, we either lose details (if the model is too smooth) or introduce

artifacts (if the model is too rough), or both, in different parts of the model. The optimal resolution scheme allows us, instead, to optimise the amount of the structural information extracted from the data while keeping estimated model errors below a consistent threshold.

4 INVERSION FOR SHEAR VELOCITY STRUCTURE

The optimal-resolution phase-velocity maps are now inverted, point by point at each knot of the model grid, for 1D, shear-velocity profiles. These profiles are then combined to form a 3D model of the crust and upper mantle. The regularisation of these 1D inversions is the same at every knot.

4.1 Removal of noisy tails of local dispersion curves

Before proceeding with the 1D inversion of the local, phase-velocity curves for shear-velocity structure, we apply additional quality checks to the dispersion curves. Even though the local curves are largely smooth, by construction, some of them have noisy "tails"—usually, the shortest-period portion of the curve having more noise than the rest of it (Fig. 20). The noisy tails are identified by evaluating the relative misfit for the edge portions of each dispersion curve (8 points for the shorter periods and 3 for the longer periods); this misfit, calculated as in Section 3 using a weakly regularized inversion, quantifies the roughness and, by inference, the error of the curve. If the misfit at any point within the tail exceeds an empirical threshold of 1.6 times the standard deviation over the full period range of the curve, then the entire tail is removed.

4.2 Gradient-search inversions

Ravenna & Lebedev (2018) showed that a well-tuned, non-linear, gradient-search inversion of phase-velocity curves yields robust 1D Earth models nearly indistinguishable from the median model given by a Bayesian, McMC inversion scheme. While a Bayesian approach still has advantages in providing a posteriori model uncertainties, we opted for a non-linear gradient-search technique, as the dataset size would have made probabilistic inversions prohibitively expensive. We do, however, perform a series of gradient inversions in each case, using variable regularization to sample the model uncertainty at different depths.

We use the same non-linear, Levenberg-Marquardt gradient search inversion algorithm that we used to isolate the noise in Section 3, with the difference that the inversions are not regularized too weakly. We choose regularization in the form of norm damping on each depth parameter, sufficient to penalize and exclude unrealistic, oscillatory V_S models. Fig. 21 illustrates how the V_S profiles with insufficient regularization (left) provide a marginally better misfit than sufficiently regularised ones (right) but are oscillatory and likely to be fitting the noise.

The algorithm computes synthetic phase velocities at each iteration directly from V_S, compressional velocity (V_P), density, and attenuation using the forward solver MINEOS (Masters et al. 2007), adapted for the travelling wave decomposition (Nolet 2008) and streamlined for speed (Lebedev et al. 2013; Ravenna & Lebedev 2018). The depth-dependent ratio between V_S and V_P is kept fixed during the inversion, equal to the values in the reference model. Density and the compressional and shear attenuation factors are fixed at the reference values, taken from PREM (Dziewonski & Anderson 1981) and AK135 (Kennett et al. 1995), respectively. The perturbations in the model, from the surface to the shallow lower mantle ($\sim$ 1300 km), are controlled using 13 triangular-shaped basis functions in the mantle and 3 boxcar-shaped ones in the crust (e.g., Bartzsch et al. 2011; Agius & Lebedev 2013). The triangular basis functions are defined by linear interpolation between neighbouring depth knots; the boxcar ones represent constant-velocity layers. The depth of the Moho and two intra-crustal discontinuities are additional inversion parameters.

4.3 Reference model

We used a three-layered reference crustal model of the region based on CRUST 2.0 (Bassin et al. 2000) and previous studies in the area (Landes et al. 2000; Tomlinson et al. 2006; Davis et al. 2012; Licciardi et al. 2014), with a reference Moho depth of 30 km. The reference model for the mantle is a modified version of AK135 (Kennett et al. 1995), recomputed at 50 s and characterized by constant shear velocities (4.45 km/s) from the Moho down to 190 km depth and linearly increasing shear velocities below. The reference period of the V_S model is 50 s, approximately in the middle of the period range of the data, in the logarithmic sense. This minimizes errors from

unknown variations in the attenuation structure of the mantle (Lebedev & van der Hilst 2008). The density for the reference model was taken from PREM (Dziewonski & Anderson 1981), as AK135 presents a questionable sharp density increase with radius in the uppermost mantle. We performed extensive tests, however, and established that this modification does not substantially change our results. The reference model is plotted in Figs. 21, 22 with a dashed black line.

4.4 Regularization of the inversion

Suitable regularization of the inversion for V_S profiles was determined in a series of tests. The regularization is by means of norm damping on the inversion parameters and is the same at all the grid knots. Instead of one inversion per location, we performed a series of 900 inversions at each point that produced a bundle of possible models, all fitting the data approximately equally well. The relative damping for each basis function is allowed to vary randomly within a range of values. This produces a set of 50 models for each knot at each damping level ($d_{\mathrm{j}=0,1,2,\ldots,49}$) as, for example, in each of the panels in the middle row in Fig. 21. A global damping factor, constant for all basis function, then multiplies the relative factors in each inversion ($D \times d_{\mathrm{j}=0,1,2,\ldots,49}$, with $D = 1, 2, 3, \ldots, 18$). The relative damping d_{j} is randomly changed at different depths (each basis function has a different value of d_{j}) in order to minimize the potential bias due to a subjective choice of regularization. The overall damping factor (D) is used to obtain differently regularized bundles of models.

Fig. 21 shows 4 differently damped (D) sets of models, with each set including 50 models with randomly varying relative damping (d_{j}). In order to cover a wide range of potentially suitable regularization parameters, 18 differently damped (varying D) models were computed that produced $4328 \times 50 \times 18$ independent, non-linear, gradient-search inversions. This allowed us to identify an optimal damping strength (D) for our final shear velocity model. We note that the optimal damping strength for the shear-velocity inversion is equal at every knot of the grid and does not vary as in the composite, phase-velocity maps (Section 3).

4.5 V_S model construction

In Fig. 22 we show sets of inversions at five different locations for a defined choice of damping factor D. The optimal damping factor is chosen according to the following criteria. We do not wish to introduce unnecessary complexity in our model by overfitting the data; if a simpler model fits the data equally well (according to the general principle of parsimony, also known as Occam's razor (e.g., Constable et al. 1987; Bodin et al. 2016)) then we choose this over a more complex model.

The accepted V_S profiles are then resampled at a 0.1 km step, and the final V_S maps are constructed, for each global-damping level. The 1D profile at a point is computed as an average over the bundle of the 50 models obtained at this knot. Maps of our final velocity model are plotted in Figs. 23, A2. A comparison of the shear-velocity maps computed with different damping levels is in Fig. A3, where we show how, for a reasonable choice of damping, the maps exhibit similar features, although the amplitudes are not preserved. One-dimensional inversions with stronger damping consistently produce smoother 3D models.

4.6 3D shear velocity structure beneath Ireland and Britain

Figs. 23 and 24 show the S-wave velocity structure of the crust and upper mantle beneath Ireland and Britain and the map of the Moho depth, resolved by our tomography. The Moho map constrained by our surface-wave inversion (Fig. 24) shows general agreement with published results from wide-angle profiles and receiver functions (e.g., Landes et al. 2007; Kelly et al. 2007; Davis et al. 2012; Licciardi et al. 2014). Our results and the ones obtained by Licciardi et al. (2014) using teleseismic P-wave receiver functions both display a thinning of the crust from SW to NE in Ireland, but also present some differences. Constrained by more data, our Moho depth map is less smooth than that of Licciardi et al. (2014). The broad agreement of our results on the topography of the Moho with published independent evidence confirms that the V_S distributions in our models are not biased substantially by trade-offs with the crustal thickness.

The S-wave velocity model (Fig. 23) offers important new insights into the structure and evolution of the Ireland-Britain region. A robust, low-velocity anomaly beneath the Irish Sea and its

surroundings persists in the models from $\sim$ 50 to at least 140 km depth, indicating an anomalously thin lithosphere, underlain by warm asthenosphere. The model represents a substantial addition to the geophysical evidence on the lithospheric structure and evolution of the Ireland-Britain region.

5 DISCUSSION

5.1 Optimal resolution tomography

Our optimal resolution tomography is developed to solve the problem as posed by Backus & Gilbert (1970): it determines localized averages over optimal resolving lengths at every point. The optimal resolving (averaging) lengths are the shortest such that the model error is below a specified level. The averaging length is a function of the global smoothing of the model, with the optimal level chosen using many inversions with different smoothing strengths. The key element of the method is the direct evaluation of the posterior model error of the phase velocity maps, which can come from both random and systematic errors in the data, as well as from the incompleteness of data sampling. In other words, the optimal resolution is chosen as the highest achievable subject to the errors.

Importantly, the estimated error decreases monotonically with the increasing smoothing of the 2D tomographic maps. This allows us to use a threshold on the errors to determine the optimal resolution. Optimal resolution does not scale with the density of the data coverage: due to systematic data errors, some of the best-sampled locations display relatively low lateral resolution. At long periods, the resolution is generally lower due to the greater wavelengths of the waves but, typically, so is the sampled deep heterogeneity, which mitigates the negative effect of the resolution decrease with increasing period. We observe that the resolution of the maps (Fig. 16) at different periods is similar, confirming the stability of our method in terms of the weak dependence of the resolving lengths on the period of the phase-velocity maps. In view of the very high heterogeneity of seismometers' distribution on the Earth's surface and the persistent and ubiquitous systematic errors in the seismic data, the new, optimal resolution tomography method can benefit studies in many regions and at different scales.

5.2 Lithospheric heterogeneity

Our tomography reveals substantial, previously unknown lithospheric heterogeneity in the area, offering exciting new insights into the structure and evolution of the Ireland-Britain region. Constrained by abundant, newly available data, our model demonstrates that the assumption of a constant lithospheric thickness across Britain and Ireland, used in the past, is not valid. At the upper-crustal depths, Ireland and Scotland exhibit similar seismic velocities, both showing clear boundaries between low-velocity sediments in basins and the high-velocity crystalline crust elsewhere, as one would expect from the continuity of the geological terrane boundaries across the Irish Sea. These contrasts are already evident in the phase-velocity maps, e.g., at 14 s (Fig. 17). In the deeper crust and lithospheric mantle, most of Ireland is surprisingly different from Scotland, showing substantially higher velocities. Our phase-velocity maps over the Irish landmass are in agreement with the earlier results from Polat et al. (2012) but provide much greater regional detail. Both studies show high velocities in the northwestern and lower velocities in the eastern part of the island. Our model also shows good agreement with the results of active seismic surveys (e.g., Landes et al. 2005), including on the Moho depth (Section 5.3).

The thinning of the lithosphere beneath the circum-Irish Sea region, evident from the model (e.g., Figs. 17, 23), matches the area of the Paleogene uplift and volcanism and offers important evidence on their mechanisms. The high velocity anomaly in west-central and east-central Ireland reveals a surprisingly thick lithosphere and may indicate the incorporation of previously unknown Precambrian continental blocks into the Irish landmass during the Caledonian Orogeny. This interpretation is consistent with the observations of Precambrian rocks in the north and west of Ireland (Chew & Stillman 2009). Detailed investigation of these features in the context of the available geological evidence will be the subject of forthcoming publications.

5.3 The Moho depth

The large-scale variations of the Moho depth across the region are generally consistent across different recent studies (Landes et al. 2005; Tomlinson et al. 2006; Kelly et al. 2007; Davis et al. 2012; Licciardi et al. 2014, 2020). Ireland and Britain generally show thicker crust (32–35 km) in

Avalonia, south of the Iapetus Suture Zone (ISZ) (Fig. 23), and thinner crust north of the ISZ, down to as thin as 26 km in northwest Scotland and Ireland. Davis et al. (2012) estimated Moho map for several profiles using receiver functions and found that the crustal thickness across Ireland and Britain varies between 24 and 36 km, with thicker crust beneath north Wales and central Scotland and thinner crust beneath northwest Scotland and northwest Ireland. Landes et al. (2005) used data from 11 seismic refraction profiles onshore and offshore to investigate the crustal velocity structure of Ireland and surrounding seas and reported a crustal thickness varying from 28.5 to 32 km in Ireland. Kelly et al. (2007) compiled a regional model for crustal seismic P-wave velocities for NW Europe, from wide-angle reflection and refraction profiles and found, for Britain and Ireland, Moho depths similar to existing crustal thickness maps, including estimates of the uncertainties for the crustal thickness and velocities. Tomlinson et al. (2006) computed the crustal thickness from the teleseismic receiver functions analysis. The results from this work on the Moho depth broadly agree with the results from seismic reflection and refraction profiles (according to the authors, ± 2 km). Some receiver functions measurements from the stations close to the ISZ show a difference with other results of up to 5 km, but apart from this the Moho depths yielded by studies using different data types are generally consistent. One notable exception is the inconsistency regarding the crustal thickness in Wales, seen between our results and, e.g., Tomlinson et al. (2006), Maguire et al. (2011), Davis et al. (2012), with the Moho shallower in our model.

6 CONCLUSIONS

The resolving length of a model varies spatially and depends on the data sampling and errors in the data (Backus & Gilbert 1970). If the posterior model errors can be evaluated directly, then the optimal width of the averaging kernels—defined as the smallest width such that the error of the local average is below a specified threshold—can be determined at every point, and optimal resolution seismic tomography can be performed.

The optimal resolution tomography scheme described in this paper relies on the direct estimation of the model error at each point, which we implemented specifically for phase-velocity tomography. It then utilizes the fact that the errors of surface-wave phase-velocity maps increase

nearly monotonically with the increasing map roughness. Thanks to this, an error threshold can be used as an effective means of determining the optimal resolving length of the tomography at every point—that is, the optimal resolution.

The validity of the scheme depends on the actual resolution of the phase-velocity maps being similar at different periods, which we observe to be the case. At the longest periods, the resolution decreases, due to the increase of the wavelength of the waves, but this is mitigated by the decrease of the heterogeneity at the long periods.

Our large, new surface-wave dataset from Ireland and Britain is characterised by unprecedentedly dense but highly heterogeneous data sampling. It also presents substantial systematic errors, evidenced by the best-sampled areas not always displaying the highest resolution. The application of the method to this dataset has provided a rigorous test for it. Our optimal-resolution tomography reveals strong, previously unknown lithospheric heterogeneity beneath Ireland and Britain and offers new insights into the structure and evolution of the region.

DATA AVAILABILITY The waveform data from the Irish National Seismic Network (INSN) (https://doi.org/10.7914/SN/EI) are available from the international data centres (http://geofon.gfz-potsdam.de/fdsnws/dataselect/1/, http://service.iris.edu/fdsnws). The waveform data from Ireland Array will be available from the international data centres after a 1-year embargo period after the end of the project in 2021-2022. The waveform data from the Great Britain Seismograph Network are available from the international data centres (http://service.iris.edu/fdsnws, http://www.orfeus-eu.org/fdsnws).

REFERENCES

Abbe, E., 1873. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung, *Archiv für mikroskopische Anatomie*, **9**(1), 413–468.

Agius, M. & Lebedev, S., 2013. Tibetan and Indian lithospheres in the upper mantle beneath Tibet: Evidence from broadband surface-wave dispersion, *Geochemistry, Geophysics, Geosystems*, **14 (10)**, 4260–4281.

Agius, M. & Lebedev, S., 2014. Shear-velocity structure, radial anisotropy and dynamics of the Tibetan crust, *Geophysical Journal International*, **199**, 1395–1415.

Aki, K. & Lee, W., 1976. Determination of three-dimensional velocity anomalies under a seismic array using first p arrival times from local earthquakes: 1. a homogeneous initial model, *Journal of Geophysical research*, **81**(23), 4381–4399.

Al-Kindi, S., White, N., Sinha, M., England, R., & Tiley, R., 2003. Crustal trace of a hot convective sheet, *Geology*, **31**(3), 207–210.

Arrowsmith, S. J., Kendall, M., White, N., VanDecar, J. C., & Booth, D. C., 2005. Seismic imaging of a hot upwelling beneath the British Isles, *Geology*, **33**(5), 345–348.

Asencio, E., Knapp, J. H., Owens, T. J., & Helffrich, G., 2003. Mapping fine-scale heterogeneities within the continental mantle lithosphere beneath Scotland: Combining active- and passive-source seismology, *Geology*, **31(6)**, 477–480.

AWE, 2020. Awe blacknest - https://www.awe.co.uk/about-us/our-locations/.

Backus, G. & Gilbert, J., 1968. The resolving power of gross earth data, *Geophys. J. Roy. Astr. Soc.,*, **16**, 169–205.

Backus, G. & Gilbert, J., 1970. Uniqueness in the inversion of inaccurate gross Earth data, *Phil. Trans. Roy. Soc. A.*, **266(1173)**, 169–205.

Bamford, D., Nunn, K., Prodehl, C., & Jacob, B., 1978. LISPB IV. Crustal structure of northern Britain, *Geophys. J. R. astr. Soc.*, **54**, 43–60.

Baptie, B., 2018. Earthquake seismology 2017/2018. open report or/18/029, british geological survey.

Barmin, M., Ritzwoller, M., & Levshin, A., 2001. A Fast and Reliable Method for Surface Wave Tomography, *Pure and Applied Geophysics*, **158**, 1351–1375.

Barton, P., 1992. LISPB revisited: a new look under the Caledonides of northern Britain, *Geophys. J. Int.*, **110**, 371–391.

Bartzsch, S., Lebedev, S., & Meier, T., 2011. Resolving the lithosphere-asthenosphere boundary with seismic Rayleigh waves, *Geophysical Journal International*, **186**, 1152–1164.

Bassin, C., Laske, G., & Masters, G., 2000. The current limits of resolution for surface-wave tomography in NorthAmerica, *Eos, Transactions, American Geophysical Union*, **81**.

Bastow, I. D., Owens, T. J., Helffrich, G., & Knapp, J. H., 2007. Spatial and temporal constraints on sources

of seismic anisotropy: Evidence from the Scottish highlands, *Geophysical Research Letters*, **34**(5).

Baykiev, E., Guerri, M., & Fullea, J., 2018. Integrating gravity and surface elevation with magnetic data: Mapping the curie temperature beneath the british isles and surrounding areas, *Frontiers in Earth Science*, **6**(October), 1–19.

Becker, T. W. & Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models, *Geochemistry, Geophysics, Geosystems*, **3**(1).

Bijwaard, H. & Spakman, W., 2000. Non-linear global P-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, **141**, 71–82.

Bijwaard, H., Spakman, W., & Engdahl, E. R., 1998. Closing the gap between regional and global travel time tomography, *Journal of Geophysical Research: Solid Earth*, **103(B12)**, 30055–30078.

Blake, T., Lebedev, S., O'Reilly, B., Agostinetti, N., Agius, M., & Schaeffer, A., 2012. An unusual occurrence of a moderately sized earthquake (ml 4.2) on the irish continental shelf and passive margin, *AGUFM*, pp. S53A–2476.

Bodin, T. & Sambridge, M., 2009. Seismic tomography with the reversible jump algorithm, *Geophysical Journal International*, **178**, 1411–1436.

Bodin, T., Sambridge, M., Tkalcic, H., Arroucau, P., Gallagher, K., & Rawlinson, N., 2012. Transdimensional inversion of receiver functions and surface-wave dispersion, *Journal of Geophysical Research*, **117**.

Bodin, T., Leiva, J., Romanowicz, B., Maupin, V., & Yuan, H., 2016. Imaging Anisotropic Layering with Bayesian Inversion of Multiple Data Types, *Geophysical Journal International*.

Bonadio, R., Geissler, W. H., Lebedev, S., Fullea, J., Ravenna, M., Celli, N. L., Jokat, W., Jegen, M., Sens-Schönfelder, C., & Baba, K., 2018. Hot upper mantle beneath the Tristan da Cunha hotspot from probabilistic Rayleigh-wave inversion and petrological modeling, *Geochemistry, Geophysics, Geosystems*, **19**, 1412–1428.

Boschi, L., 2003. Measures of resolution in global body wave tomography, *Geophysical Research Letters*, **30(19)**, 25567–25594.

Boschi, L. & Dziewonski, A., 1999. High-and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *JGR Solid Earth*, **104(B11)**, 25567–25594.

Boschi, L., Weemstra, C., Verbeke, J., Ekström, G., Zunino, A., & Giardini, D., 2013. On measuring surface-wave phase velocity from station-station cross-correlation of ambient signal, *Geophys J. Int.*, **192**, 346–358.

Bott, M., Long, R., Green, A., Lewis, A., Sinha, M., & Stevenson, D., 1985. Crustal structure south of the Iapetus suture beneath northern England, *Nature*, **314**, 724–727.

Brown, C. & Whelan, J. P., 1955. Terrane boundaries in Ireland inferred from the Irish Magnetotelluric Profile and other geophysical data, *Journal of the Geological Society of London*, **152**, 523–534.

Celli, N., Lebedev, S., Schaeffer, A., Ravenna, M., & Gaina, C., 2020. The upper mantle beneath the South Atlantic Ocean, South America and Africa from waveform tomography with massive data sets, *Geophys. J. Int.*, **221**, 178–204.

Chadwick, R. & Pharaoh, T., 1998. The seismic reflection Moho beneath the United Kingdom and adjacent areas, *Tectonophysics*, **299**, 255–279.

Chew, D. M. & Stillman, C. J., 2009. *Late Caledonian orogeny and magmatism*, Dunedin Academic Press ltd.

Chou, C. & Booker, J., 1979. A Backus–Gilbert approach to inversion of travel time data for three-dimensional velocity structure, *Geophys. J. Royal Astr. Soc.*, **59**, 325–344.

Cogné, N., Doepke, D., Chew, D., Stuart, F. M., & Mark, C., 2016. Measuring plume-related exhumation of the British Isles in Early Cenozoic times, *Earth and Planetary Science Letters*, **456**, 1–15.

Constable, S., Parker, R., & Constable, C., 1987. Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data., *GEOPHYSICS*, **52(3)**, 289–300.

Curtis, A. & Snieder, R., 1997. Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, **62**, 1524–1532.

Dahlen, F. & Zhou, Y., 2006. Surface-wave group-delay and attenuation kernels, *Geophysical Journal International*, **165(2)**, 545–554.

Dahlen, F. A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press.

Darbyshire, F. A. & Lebedev, S., 2009. Rayleigh wave phase-velocity heterogeneity and multilayered azimuthal anisotropy of the Superior Craton, Ontario, *Geophys. J. Int.*, **176**, 215–234.

Davis, M., White, N., Priestley, K., Baptie, B., & Tilmann, F., 2012. Crustal structure of the British Isles and its epeirogenic consequences, *Geophysical Journal International*, **190**, 705–725.

De Laat, J., Celli, N., & Lebedev, S., 2019. Imaging the North Atlantic: deployment of the ocean-bottom seismometers of the SEA-SEIS project and quality control of global and regional network data, *Geophysical Research Abstracts*, **21**.

de Vos, D., Paulssen, H., & Fichtner, A., 2013. Finite-frequency sensitivity kernels for two-station surface wave measurements, *Geophysical Journal International*, **194**(2), 1042–1049.

Debayle, E. & Sambridge, M., 2004. Inversion of massive surface-wave data sets: Model construction and resolution assessment, *Journal of Geophysical Research*, **109**.

Demmerle, J., Wegel, E., Schermelleh, L., & Dobbie, I., 2015. Assessing resolution in super-resolution imaging, *Methods*, **88**.

den Dekker, A. & van den Bos, A., 1997. Resolution: a survey, *J. Opt. Soc. Am. A*, **14(3)**.

Deschamps, F., Lebedev, S., Meier, T., & Trampert, J., 2008. Azimuthal anisotropy of Rayleigh-wave phase velocities in the east-central United States, *Geophysical Journal International*, **173**, 827–843.

Di Leo, J., Bastow, I., & Helffrich, G., 2009. Nature of the Moho beneath the Scottish Highlands from a

receiver function perspective, *Tectonophysics*, **474**, 214–222.

Do, V., O'Reilly, B., Landes, M., & Readman, P., 2006. Shear-wave splitting observations across southwest Ireland, *Geophys. Res. Lett.*, **33(L03309)**, 1–23.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model, *Physics of the Earth and Planetary Interiors*, **25**, 297–356.

Dziewonski, A. M., Hager, B. H., & O'Connell, R. J., 1977. Large-scale heterogeneities in the lower mantle, *Journal of Geophysical Research*, **82**(2), 239–255.

Dziewonski, A. M., Chou, T.-A., & Woodhouse, J. H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *Journal of Geophysical Research*, **86**, 2825–2852.

Edwards, J. & Blundell, D., 1984. Summary of seismic refraction experiments in the English Channel, Celtic Sea and St. George's Channel, *British Geological Survey Marine Geophysics Report*, **144**.

Ekström, G., Nettles, M., & Dziewonski, A., 2012. The global CMT project 2004–2010: Centroid-moment tensors for 13017 earthquakes, *Physics of the Earth and Planetary Interiors*, **200–201**, 1–9.

El-Sharkawy, A., Meier, T., Lebedev, S., Behrmann, J. H., Hamada, M., Cristiano, L., Weidle, C., & Köhn, D., 2020. The slab puzzle of the alpine-mediterranean region: Insights from a new, high-resolution, shear wave velocity model of the upper mantle, *Geochemistry, Geophysics, Geosystems*, **21**(8), e2020GC008993.

Endrun, B., Meier, T., Lebedev, S., Bohnhoff, M., Stavrakakis, G., & Harjes, H.-P., 2008. S velocity structure and radial anisotropy in the Aegean region from surface-wave dispersion, *Geophys. J. Int.*, **174**, 593–616.

Erduran, M., Endrun, B., & Meier, T., 2008. Continental vs. oceanic lithosphere beneath the eastern Mediterranean Sea – implications from Rayleigh wave dispersion measurements, *Tectonophysics*, **457**, 42–52.

Feynman, R., Leighton, R., & Sands, M., 1963. The Feynman Lectures on Physics, *Reading, Mass: Addison-Wesley Pub. Co.*.

Fitchner, A. & van Leeuwen, T., 2015. Resolution analysis by random probing, *JGR Solid Earth*.

Freeman, B., Klemperer, S., & Hobbs, R., 1988. The deep structure of northern England and the Iapetus Suture zone from BIRPS deep seismic reflection profiles, *J. geol. Soc. Lond.*, **145**, 727–740.

Fry, B., Boschi, L., Ekstrom, G., & Giardini, D., 2008. Europe-Mediterranean tomography: high correlation between new seismic data and independent geophysical observables, *Geophys. Res. Lett.*, **35(L04301)**.

Fukao, Y., Obayashi, M., Inoue, H., & Nebai, M., 1992. Subducting slabs stagnant in the mantle, *J. Geophys. Res.*, **97**, 4809–4822.

Fullea, J., Muller, M. R., Jones, A. G., & Afonso, J. C., 2014. The lithosphere-asthenosphere system

beneath Ireland from integrated geophysical-petrological modeling II: 3D thermal and compositional structure, *Lithos*, **189**, 49–64.

Galetti, E., Curtis, A., Baptie, B., Jenkins, D., & Nicolson, H., 2016. Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry, *Geophysical Journal International*, **208**, 36–58.

GEOFON Data Centre, 1993. Geofon seismic network. deutsches geoforschungszentrum gfz., 10.14470/tr560404.

Hardwick, A., 2008. *New insights into the crustal structure of the England, Wales and Irish Seas areas from local earthquake tomography and associated seismological studies*, Ph.D. thesis, University of Leicester.

Hawkins, R., Bodin, T., Sambridge, M., Choblet, G., & Husson, L., 2019. Trans-dimensional surface reconstruction with different classes of parameterization, *Geochemistry, Geophysics, Geosystems*.

Helmholtz, H., 1874. On the theoretical limits of the optical capacity of the microscope, *Monthly Microscopical Journal*, **16**, 15–39.

Hodgson, J., 2001. *A seismic and gravity study of the Leinster Granite: SE Ireland*, Ph.D. thesis, Dublin Institute for Advanced Studies, Dublin, Ireland.

Holland, C. & Sanders, I., 2009. *The Geology of Ireland, 2nd edn*, Dunedin Academic Press, Edinburgh, Scotland.

INSN, 1993. Irish national seismic network, operated by the dublin institute for advanced studies and supported by the geological survey ireland. international federation of digital seismograph networks, doi.org/10.7914/sn/ei.

IOC, IHO, & BODC, 2003. Centenary Edition of the GEBCO Digital Atlas, `https://www.gebco.net`.

Jacob, A., Kaminski, W., Murphy, T., Phillips, W., & Prodehl, C., 1985. A crustal model for a northeast-southwest profile through Ireland, *Tectonophysics*, **113(1–2)**, 75–103.

Jones, A., Afonso, J., Fullea, J., Muller, M., & Salajegheh, F., 2013. The lithosphere–asthenosphere system beneath Ireland from integrated geophysical–petrological modeling – I: Observations, 1D and 2D modelling, *Lithos*, **189**.

Jones, S. M., White, N., Clarke, B. J., Rowley, E., & Gallagher, K., 2002. Present and past influence of the Iceland Plume on sedimentation, *Geological Society, London, Special Publications*, **196**, 13–25.

Kelly, A., England, R. W., & Maguire, P. K., 2007. A crustal seismic velocity model for the UK, Ireland and surrounding seas, *Geophysical Journal International*, **171**(3), 1172–1184.

Kennett, B., Engdahl, E., & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel-times, *Geophysical Journal International*, **122 (1)**, 108–124.

Klemperer, S. & Hobbs, R., 1991. The BIRPS Atlas: Deep Seismic Reflection Profiles Around the British Isles, *Cambridge Univ. Press.*.

Klemperer, S., Ryan, P., & Snyder, D., 1991. A deep seismic reflection transect across the Irish Caledonides, *J. Geol. Soc. Lond.*, **148**, 149–164.

Köhler, H., 1981. On abbe's theory of image formation in the microscope, *Optica Acta: International Journal of Optics*, **28**(12), 1691–1701.

Kolínskỳ, P., Bokelmann, G., & Group, A. W., 2021. On the wobbles of phase-velocity dispersion curves, *Geophysical Journal International*, **224**(3), 1477–1504.

Landes, M., Prodehl, C., Hauser, F., Jacob, A., & Vermeulen, N., 2000. VARNET-96: influence of the Variscan and Caledonian orogenies on crustal structure in SW Ireland, *Geophysical Journal International*, **140 (3)**, 660–676.

Landes, M., Ritter, J., Do, V., Readman, P., & O'Reilly, B., 2004. Passive teleseismic experiment explores the deep subsurface of southern Ireland, *Eos, Trans. - Am. Geophys. Union*, **85(36)**, 337–341.

Landes, M., Ritter, J., Readman, P., & O'Reilly, B., 2005. A review of the Irish crustal structure and signatures from the Caledonian and Variscan orogenies, *Terra Nova*, **17 (2)**, 111–120.

Landes, M., Ritter, J., O'Reilly, B., Readman, P., & Do, V., 2006. A N–S receiver function profile across the Variscides and Caledonides in SW Ireland, *Geophys. J. Int.*, **166**, 814–824.

Landes, M., Ritter, J., & Readman, P., 2007. Proto-Iceland plume caused thinning of Irish lithosphere, *Earth and Planetary Science Letters*, **255 (1)**, 32–40.

Lebedev, S. & van der Hilst, R. D., 2008. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophysical Journal International*, **173**, 505–518.

Lebedev, S., Nolet, G., & van der Hilst, R. D., 1997. The upper mantle beneath the Philippine Sea region from waveform inversions, *Geophysical Research Letters*, **24**(15), 1851–1854.

Lebedev, S., Chevrot, S., & van der Hilst, R. D., 2003. Correlation between the shear-speed structure and thickness of the mantle transition zone, *Physics of the Earth and Planetary Interiors*, **136**(1-2), 25–40.

Lebedev, S., Nolet, G., Meier, T., & van der Hilst, R. D., 2005. Automated multimode inversion of surface and S waveforms, *Geophys J. Int.*, **162**, 951–964.

Lebedev, S., Meier, T., & van der Hilst, R. D., 2006. Asthenospheric flow and origin of volcanism in the Baikal Rift area, *Earth and Planetary Science Letters*, **249**, 415–424.

Lebedev, S., Horan, C., Readman, P. W., Schaeffer, A. J., Agius, M. R., Collins, L., Hauser, F., O'Reilly, B. M., & Blake, T., 2012. Ireland Array: a new broadband seismic network targets the structure, evolution and seismicity of Ireland and surroundings, *Proceedings of EGU General Assembly*.

Lebedev, S., Adam, J. M.-C., & Meier, T., 2013. Mapping the Moho with seismic surface-waves: A review, resolution analysis, and recommended inversion strategies, *Tectonophysics*, **609**, 377–394.

Licciardi, A. & Piana Agostinetti, N., 2014. High frequency receiver functions in the Dublin Basin: application to a potential geothermal site, *Energy Procedia*, **59**, 221–226.

Licciardi, A. & Piana Agostinetti, N., 2017. Sedimentary basin exploration with receiver functions: Seis-

mic structure and anisotropy of the Dublin Basin (Ireland), *Geophysics*, **82**, KS41–KS55.

Licciardi, A., Piana Agostinetti, N., Lebedev, S., Schaeffer, A., Readman, P., & Horan, C., 2014. Moho depth and v_p/v_s in Ireland from teleseismic receiver functions analysis, *Geophys. J. Int.*, **199**, 561–579.

Licciardi, A., England, R., Piana Agostinetti, N., & Gallagher, K., 2020. Moho depth of the British Isles: a probabilistic perspective, *Geophys. J. Int.*.

Lowe, C. & Jacob, A., 1989. A north-south seismic profile across the Caledonian Suture zone in Ireland, *Tectonophysics*, **168**, 297–318.

Maguire, P., England, R., & Hardwick, A., 2011. LISPB DELTA, a lithospheric seismic profile in Britain: Analysis and interpretation of the Wales and southern England section, *Journal of the Geological Society*, **168**(1), 61–82.

Marquering, H. & Snieder, R., 1996. Surface-wave velocity structure beneath Europe, the northeastern Atlantic and western Asia from waveform inversion including surface wave mode coupling, *Geophys. J. Int.*, **127**, 238–304.

Masson, F., Jacob, A., Prodehl, C., Readman, P., Shannon, P., Schulze, A., & Enderle, U., 1998. A wide-angle seismic traverse through the Variscan of southwest Ireland, *Geophys. J. Int*, **134**, 689–705.

Masson, F., Hauser, F., & Jacob, A., 1999. The lithospheric trace of the Iapetus Suture in SW Ireland from teleseismic data, *Tectonophysics*, **302**, 83–98.

Masters, G., Barmine, M., & Kientz, S., 2007. Mineos user's manual, *Computational Infrastructure for Geodynamics, California Institute of Technology, Pasadena, `http://geodynamics.org/cig/software/mineos`.*

Mather, B. & Fullea, J., 2019. Constraining the geotherm beneath the British Isles from Bayesian inversion of Curie depth: Integrated modelling of magnetic, geothermal and seismic data, *Solid Earth Discussions*, pp. 1–20.

Mather, B., Farrell, T., & Fullea, J., 2018. Probabilistic Surface Heat Flow Estimates Assimilating Paleoclimate History: New Implications for the Thermochemical Structure of Ireland, *Journal of Geophysical Research: Solid Earth*.

McNaught, A. & Wilkinson, A., 2014. Resolution, *Compendium of chemical terminology Gold Book, version 2.3.3*, p. 1296.

Meier, T., Dietrich, K., Stockhert, B., & Harjes, H., 2004. One-dimensional models of shear wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh wave phase velocities and tectonic implications, *Geophysical Journal International*, **156 (1)**, 45–58.

Menke, W., 2012. Geophysical Data Analysis: Discrete Inverse Theory, *Academic Press*, **45**.

Menke, W., 2015. Review of the generalized least squares method, *Surveys in Geophysics*, **36**(1), 1–25.

Michelini, A., 1995. An adaptive-grid formalism for traveltime tomography, *Geophys. J. Int.*, **121**.

Möllhoff, M. & Bean, C., 2016. Seismic Noise Characterization in Proximity to Strong Microseism

Sources in the Northeast Atlantic, *Bulletin of the Seismological Society of America*, **106(2)**, 464–477.

Nicolson, H., Curtis, A., Baptie, B., & Galetti, E., 2012. Seismic interferometry and ambient noise tomography in the British Isles, *Proceedings of the Geologists' Association*, **123**, 74–86.

Nicolson, H., Curtis, A., & Baptie, B., 2014. Rayleigh wave tomography of the British Isles from ambient seismic noise, *Geophysical Journal International*, **198**, 637–655.

Nolet, G., 1985. A Seismic Tomography Program for Geological Investigations, *Joint Interpretation of Geophysical and Geological Data Applied to Lithospheric Studies*, pp. 109–126.

Nolet, G., 2008. A Breviary of Seismic Tomography, *Press, Cambridge, UK*, **324**.

O'Donnell, J. P., Daly, E. Tiberi, C., Bastow, I. D., O'Reilly, B. M., Readman, P. W., & Hauser, F., 2011. Lithosphere–asthenosphere interaction beneath Ireland from joint inversion of teleseismic P-wave delay times and GRACE gravity, *Geophys. J. Int.*.

Oldenborger, G. & Routh, P., 2009. The point-spread function measure of resolution for the 3-d electrical resistivity experiment, *Geophysical Journal International*, **176(2)**, 405–414.

O'Reilly, B., Hauser, F., Jacob, A., & Shannon, P., 1996. The lithosphere below the Rockall Trough: wide-angle seismic evidence for extensive serpentinisation, *Tectonophysics*, **255**, 1–23.

O'Reilly, B. M., Hauser, F., & Readman, P. W., 2010. The fine-scale structure of upper continental lithosphere from seismic waveform methods: Insights into Phanerozoic crustal formation processes, *Geophysical Journal International*, **180**(1), 101–124.

O'Reilly, B. M., Hauser, F., & Readman, P. W., 2012. The fine-scale seismic structure of the upper lithosphere within accreted Caledonian lithosphere: Implications for the origins of the 'Newer Granites', *Journal of the Geological Society*, **169**(5), 561–573.

Paige, C. C. & Saunders, M. A., 1987. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares, *ACM Transactions on Mathematical Software (TOMS)*, **9**, 43–71.

Parker, R., 1994. Geophysical Inverse Theory, *Princeton University Press, Princeton, NJ,*, p. 386.

Pawlak, A., Eaton, D. W., Darbyshire, F., Lebedev, S., & Bastow, I. D., 2012. Crustal anisotropy beneath hudson bay from ambient noise tomography: Evidence for post-orogenic lower-crustal flow?, *Journal of Geophysical Research: Solid Earth*, **117**(B8).

Piana Agostinetti, N. & Licciardi, A., 2015. SIM-CRUST: seismic imaging and monitoring of the upper crust: exploring the potential low-enthalpy geothermal resources of Ireland, *submitted to Proceeding of the 2015 World Geothermal Congress.*.

Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, **181(2)**, 858–872.

Piana Agostinetti, N., Giacomuzzi, G., & Malinverno, A., 2015. Local 3D earthquake tomography by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, **201(3)**, 1598–1617.

Pijpers, F. & Thompson, M., 1994. The sola method for helioseismic inversion, *Astronomy and Astro-*

physics, **281**, 231–240.

Polat, G., Lebedev, S., Readman, P. W., O'Reilly, B. M., & Hauser, F., 2012. Anisotropic Rayleigh-wave tomography of Ireland's crust: Implications for crustal accretion and evolution within the Caledonian Orogen, *Geophysical Research Letters*, **39 (L04302)**.

Rao, C., Jones, A., & Moorkamp, M., 2007. The geometry of the Iapetus Suture Zone in central Ireland deduced from a magnetotelluric study, *Phys. Earth planet. Inter.*, **161(1-2)**, 134–141.

Ravenna, M. & Lebedev, S., 2018. Bayesian inversion of surface-wave data for radial and azimuthal shear-wave anisotropy, with applications to central Mongolia and west-central Italy, *Geophysical Journal International*, **213**, 278–300.

Ravenna, M., Lebedev, S., Fullea, J., & Adam., J. M.-C., 2018. Shear-wave velocity structure of southern Africa's lithosphere: Variations in the thickness and composition of cratons and their effect on topography, *Geochem. Geophys. Geosyst.*, **19**, 1499–1518.

Rawlinson, N. & Kennett, B., 2004. Rapid estimation of relative and absolute delay times across a network by Adaptive Stacking, *Geophysical Journal International*, **157(1)**, 332–340.

Rawlinson, N. & Spakman, W., 2016. On the use of sensitivity tests in seismic tomography, *Geophysical Journal International*, **205**(2), 1221–1243.

Rawlinson, N., Pozgay, S., & Fishwick, S., 2010. Seismic tomography: A window into deep Earth, *Physics of the Earth and Planetary Interiors*.

Rayleigh, J., 1896. On the theory of optical images, with special reference to the microscope., *Philosophical Magazine*, **XLII**, 167–195.

Readman, P., O'Reilly, B., & Murphy, T., 1997. Gravity gradients and upper-crustal tectonic fabrics, Ireland, *J. Geol. Soc. Lond.*, **154**, 817–828.

Rickers, F., Fichtner, A., & Trampert, J., 2013. The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: Evidence from full-waveform inversion, *Earth and Planetary Science Letters*, **367**, 39–51.

Ritzwoller, M., Shapiro, N., Barmin, M., & Levshin, A., 2002. Global surface wave diffraction tomography, *Journal of Geophysical Research, Solid Earth*.

Ronchi, V., 1961. Resolving power of calculated and detected images, *J. Opt. Soc. Am.*, **51**, 458–460.

Sambridge, M. & Faletič, R., 2003. Adaptive whole Earth tomography, *Geochemistry, Geophysics, Geosystems*, **4**.

Sambridge, M. & Rawlinson, N., 2005. Seismic tomography with irregular meshes, *Levander, A., Nolet, G. (Eds.), Seismic Earth: Array Analysis of Broadband Seismograms, vol. 157. American Geophysical Union*, pp. 49–65.

Sambridge, M., Braun, J., & McQueen, H., 1995. Geophysical parameterization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.*, **122**, 837–857.

Sato, Y., 1955. Analysis of Dispersed Surface Waves by means of Fourier Transform I., *Bulletin of the Earthquake Research Institute*, **33**, 33–47.

Schaeffer, A., Lebedev, S., & Becker, T., 2016. Azimuthal seismic anisotropy in the Earth's upper mantle and the thickness of tectonic plates, *Geophysical Journal International*, **207**, 901–933.

Schaeffer, A. J. & Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone, *Geophysical Journal International*, **194**, 417–449.

Schaeffer, A. J. & Lebedev, S., 2015. Global heterogeneity of the lithosphere and underlying mantle: A seismological appraisal based on multimode surface-wave dispersion analysis, shear-velocity tomography, and tectonic regionalization, *Invited Review in: The Earth's Heterogeneous Mantle, A. Khan and F. Deschamps (eds.) - Springer Geophysics*, pp. 3–46.

Schivardi, R. & Morelli, A., 2009. Surface wave tomography in the European and Mediterranean region, *Geophys. J. Int.*, **177**, 1050–1066.

Shaw Champion, M., White, N., Jones, S., & Priestley, K., 2006. Crustal velocity structure of the British Isles: a comparison of receiver, *Geophys. J. Int.*, **166**, 795–813.

Sheriff, R. & Geldart, L., 1995. Exploration Seismology, *Cambridge University Press*.

Smith, M. L. & Dahlen, F. A., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *Journal of Geophysical Research*, **78**, 3321–3333.

Snyder, D. & Flack, C., 1990. A Caledonian age for reflectors within the mantle lithosphere north and west of Scotland, *Tectonics*, **9**, 903–922.

Soomro, R., Weidle, C., Cristiano, L., Lebedev, S., & Meier., T., 2016. Phase velocities of Rayleigh and Love waves in central and northern Europe from automated, broadband, inter-station measurements, *Geophysical Journal International*, **204 (1)**, 517–534.

Tarantola, A. & Nercessian, A., 1984. Three-dimensional inversion without blocks, *Geophys. J. Royal Astr. Soc.*, **76**, 299–306.

Tiley, R., White, N., & Al-Kindi, S., 2004. Linking Paleogene denudation and magmatic underplating beneath the British Isles, *Geological Magazine*, **141**(3), 345–351.

Tomlinson, J. P., Denton, P., Maguire, P. K., & Booth, D. C., 2006. Analysis of the crustal velocity structure of the British Isles using teleseismic receiver functions, *Geophysical Journal International*, **167**(1), 223–237.

Trampert, J. & van Heijst, H. J., 2002. Global azimuthal anisotropy in the transition zone, *Science*, **296**(5571), 1297–1299.

van der Hilst, R., Widiyantoro, S., & Engdahl, E., 1997. Evidence for deep mantle circulation from global tomography, *Nature*, **386(6625)**, 578–584.

Wawerzinek, B., Ritter, J. R., Jordan, M., & Landes, M., 2008. An upper-mantle upwelling underneath Ireland revealed from non-linear tomography, *Geophys. J. Int.*, **175**, 253–268.

Weidle, C., Soomro, R., Cristiano, L., & Meier, T., 2013. Identification of response and timing issues at permanent European broadband stations from automated data analysis, *Adv. Geosci.*, **36**, 21–25.

Wessel, P., Luis, J., Scharroo, R., Wobbe, F., Smith, W., & Tian, D., 2019. The generic mapping tools version 6, *Geochemistry, Geophysics, Geosystems*, **20 (11)**, 5556–5564.

White, N. & Lovell, B., 1997. Measuring the pulse of a plume with the sedimentary record, *Nature*, **387**(6636), 888–891.

Yanovskaya, T., 2005. Development of methods for surface wave tomography based on the backus-gilbert approach, *Selected Papers From Volume 32 of Vychislitel'naya Seysmologiya*, **7**, 5–16.

Zaroli, C., 2016. Global seismic tomography using Backus-Gilbert inversion, *Geophysical Journal International*, **207(2)**, 876–888.

Zaroli, C., 2019. Seismic tomography using parameter-free Backus–Gilbert inversion, *Geophysical Journal International*, **218**, 619–630.

Zaroli, C., Koelemeijer, P., & Lambotte, S., 2017. Toward seeing the Earth's interior through unbiased tomographic lenses, *Geophys. Res. Lett.*, **44**, 11399–11408.

Zhang, X., Paulssen, H., Lebedev, S., & Meier, T., 2009. 3d shear velocity structure beneath the gulf of california from rayleigh wave dispersion, *Earth and Planetary Science Letters*, **279**(3-4), 255–262.

Zhao, D., 2004. Global tomographic images of mantle plumes and subducting slabs: Insight into deep Earth dynamics, *Physics of the Earth and Planetary Interiors*, **146**(1-2), 3–34.

ACKNOWLEDGMENTS

We thank Nick Rawlinson for valuable comments on an early draft of the manuscript and the Editor, Andrea Morelli, and two anonymous reviewers for constructive suggestions that helped us to improve the paper. We are grateful to the operators of the networks and arrays in Ireland and Britain that produced the data used in this study: Ireland Array (Lebedev et al. 2012), the United Kingdom network operated by the British Geological Survey (Baptie 2018), INSN: the Irish National Seismic Network (INSN 1993; Blake et al. 2012), the Dublin Basin temporary network (Licciardi & Piana Agostinetti 2014), the ISLE and ISUME projects (Landes et al. 2004; Do et al. 2006; Landes et al. 2007; Wawerzinek et al. 2008; O'Donnell et al. 2011; Polat et al. 2012), the WAVEOBS network (Möllhoff & Bean 2016), the SIM-CRUST project (Piana Agostinetti & Licciardi 2015), the Blacknest network (AWE 2020). The figures were created with the Generic Mapping Tools (GMT, Wessel et al. 2019). ISLE was funded by a Basic Research Grant SC/01/155 from Enter-

prise Ireland, and ISUME by a Research Frontiers Programme Research Grant 07/RFP/GEOF758 from Science Foundation Ireland. This work was supported by research grants from Science Foundation Ireland (SFI) under grant number 13/CDA/2192 and grant number 13/RC/2092, co-funded under the European Regional Development Fund and by iCRAG industry partners. Additional support came from grant 16/IA/4598, cofunded by the Science Foundation Ireland, the Geological Survey of Ireland and the Marine Institute.

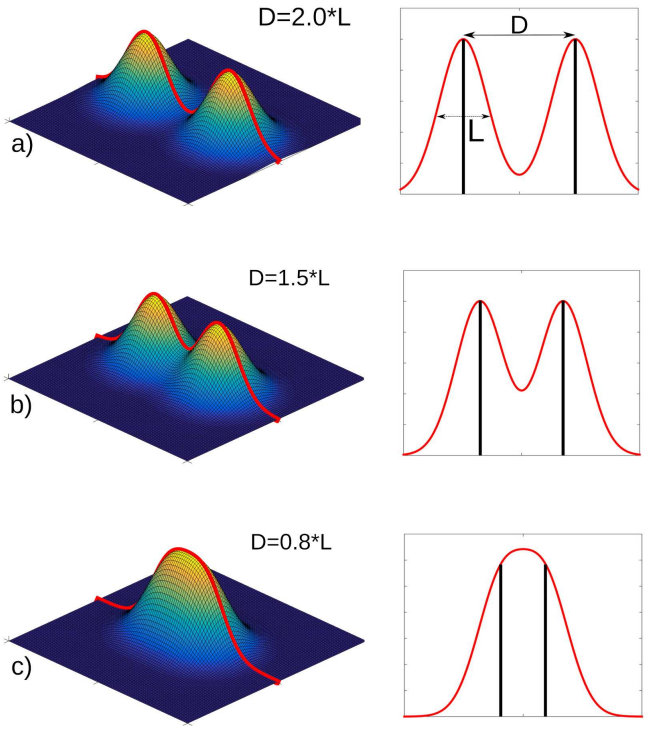

Figure 1. Schematic illustration of the averaging kernel and the resolving length. The resolving (averaging) length L is defined as the full width of the local averaging kernel at its half maximum. Resolving length equals the distance D between two point anomalies above which they can be distinguished on the image as separate.

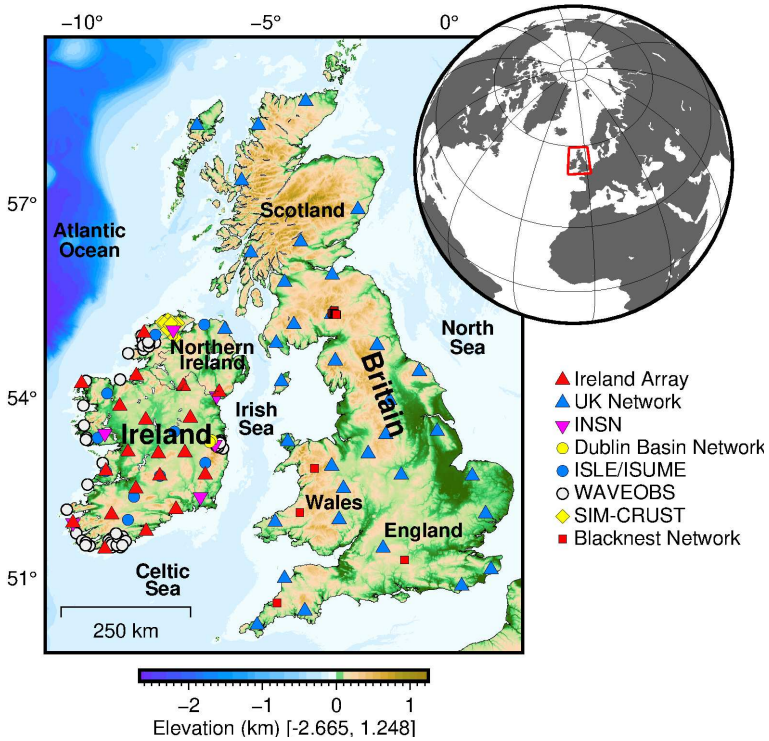

Figure 2. The seismic stations in Ireland and Britain used in this study belong to Ireland Array (Lebedev et al. 2012), the United Kingdom network operated by the British Geological Survey (e.g., Baptie 2018),the Irish National Seismic Network (INSN) (INSN 1993; Blake et al. 2012), the Dublin Basin temporary network (Licciardi & Piana Agostinetti 2014), the ISLE and ISUME projects (Landes et al. 2004; Do et al. 2006; Landes et al. 2007; Wawerzinek et al. 2008; O'Donnell et al. 2011; Polat et al. 2012), the WAVEOBS project (Möllhoff & Bean 2016), the SIM-CRUST project (Piana Agostinetti & Licciardi 2015), and the Blacknest Array (AWE 2020). Topography and bathymetry are from the GEBCO dataset (IOC et al. 2003).

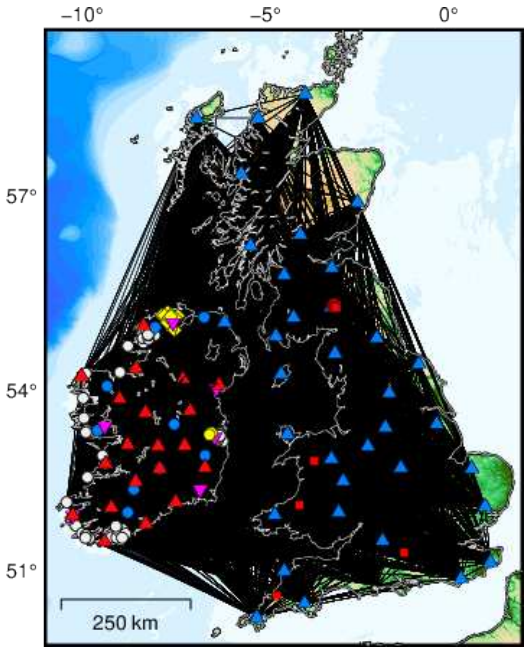

Figure 3. Station locations (as in Fig. 2) and the interstation path coverage (black lines) yielded by our measurements. The coverage in Ireland is denser than in Britain due to data availability.

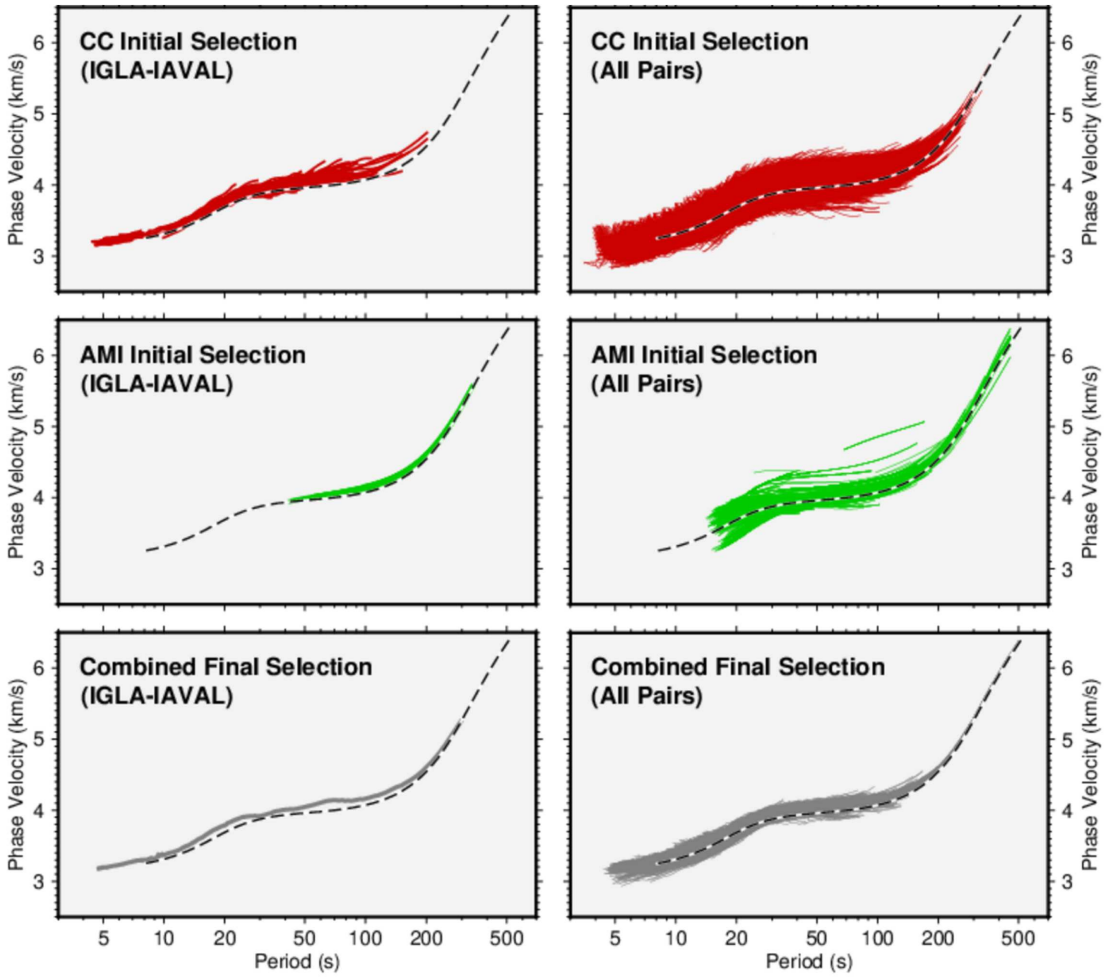

Figure 4. Phase-velocity curves measured with the two-station cross-correlation and waveform inversion. Left: the single-event, two-station dispersion measurements for the pair IGLA–IAVAL. Right: the measurements for all the station pairs. Top panels: cross correlation measurements. Middle: waveform inversion measurements. Bottom: the final phase-velocity curves computed as averages over all the measurements from the two methods for the station pair. Only the portions of the curves with at least 10 single-event measurements at each period are included.

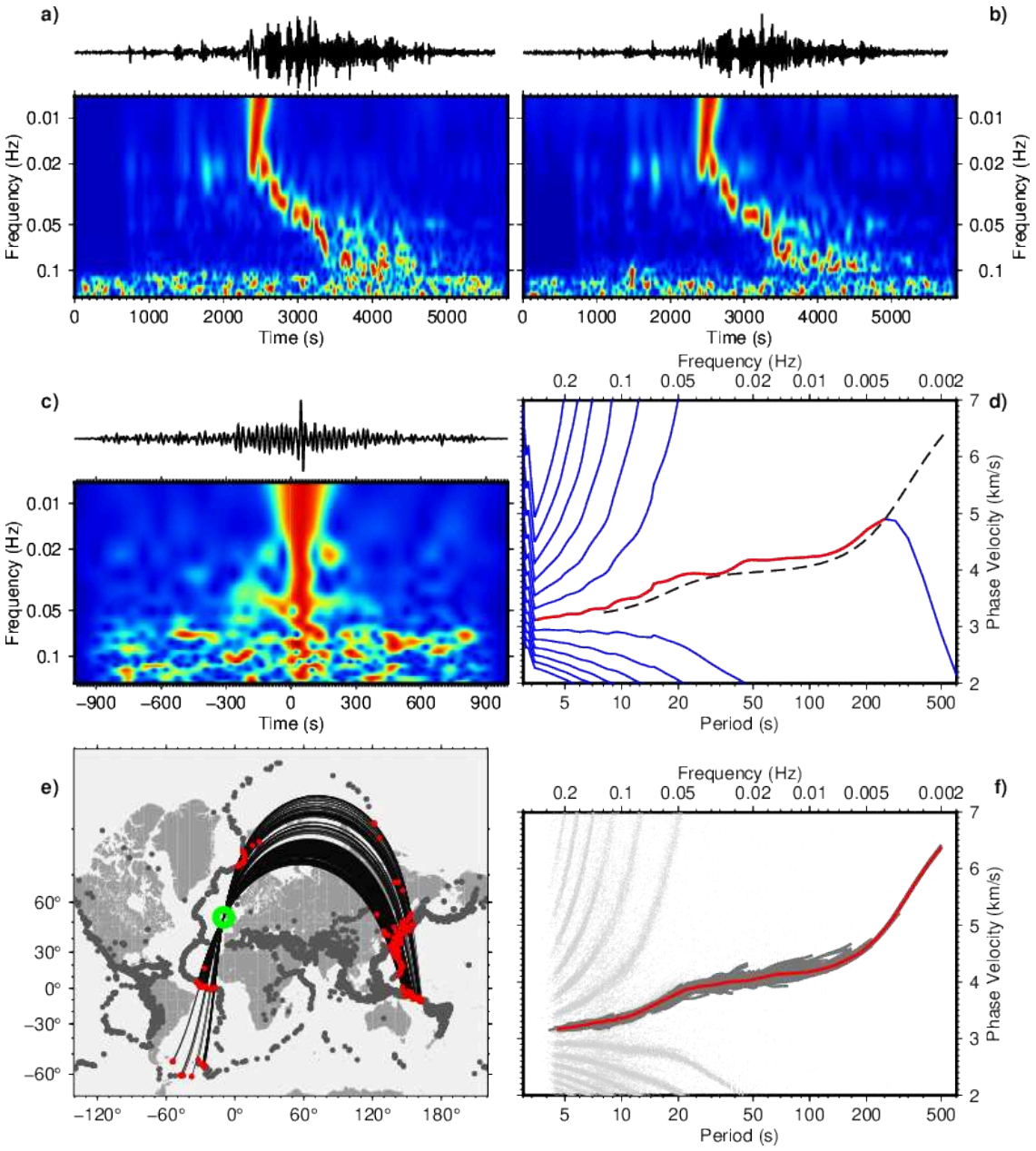

Figure 5. Example of an interstation, Rayleigh-wave, phase-velocity measurement for the station pair IGLA–IAVAL. Both IGLA and VAL are INSN stations in western Ireland, and IAVAL was an Ireland Array station collocated with VAL and operated before a broadband INSN instrument was installed at the site. Panels (a) and (b) show the recorded seismograms and the time-frequency representations of their waveforms. c): the cross-correlation signal and its time-frequency representation. d): phase-velocity curves measured in the frequency domain by unwrapping the phase of the cross-correlation function (Section 2.1). Alternative curves resulting from the 2π ambiguity are plotted in blue. The accepted segment is shown in red. Dashed line: AK135 (Kennett et al. 1995). e): all the events used in this study (grey), the events used for this pair of stations (red) and the great circle paths between these events and the stations. The locations of the two stations are within the green circle. f): the accepted one-event, phase-velocity measurements (dark grey lines) and the final measurement for the station pair (red), computed by averaging over all cross-correlation and waveform inversion measurements. The light grey dots indicate the branches of measurements affected by the 2π ambiguity, relative to the accepted one-event, phase-velocity measurements shown with dark grey lines. The event that yielded the data in a)–d) is an Mw=7.32 earthquake located at 38.56N, 142.78E, at a depth of 14.1 km. The station IGLA (at 53.42N, 9.38W) is 84.92° away from the event, with a back azimuth 21.56°. The station IAVAL (51.94N, 10.24W) is 86.49° away from the event, with a back azimuth 20.87°. The interstation distance is 174.5 km.

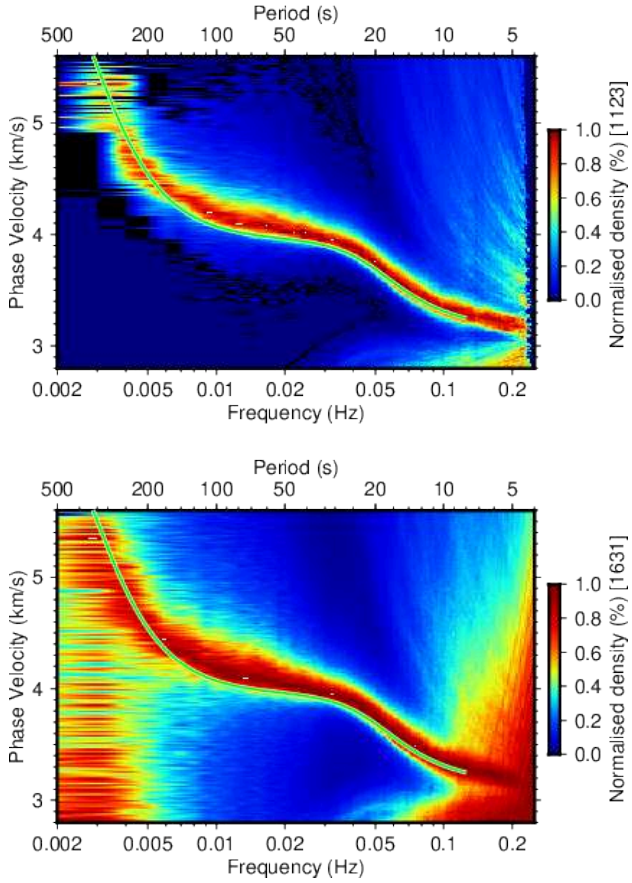

Figure 6. Determination of a reference dispersion curve. Bottom: The stack of all initial phase-velocity measurements in the region, with no selection applied, already shows a fairly accurate, region-average phase-velocity curve (Section 2.3). The stack is computed using all branches of possible phase-velocity curves, including those affected by the 2π ambiguity, for all pairs of stations. The branches that do not represent the real Earth structure tend to cancel out in the stack. Top: the stack obtained via a loose selection of preliminary measurements (Section 2.3), shows an improvement compared to the stack obtained from all measurements (bottom) and yields an accurate reference model for the definitive, one-by-one phase-velocity measurement selection. Both density plots are normalized to the maximum at each frequency. Green lines: AK135 (Kennett et al. 1995).

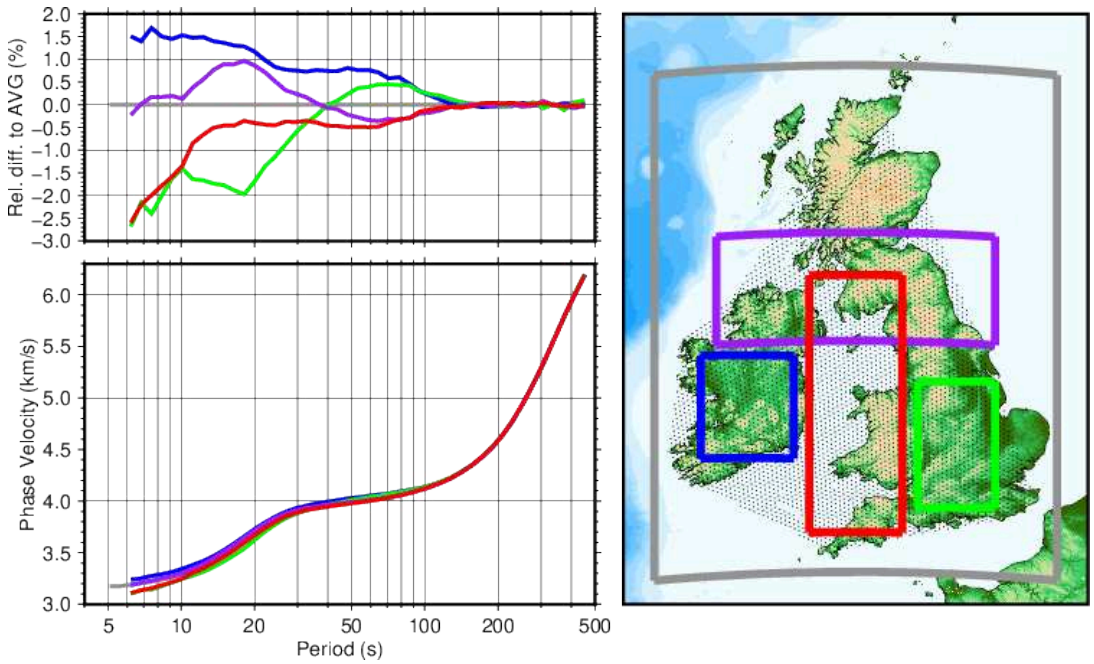

Figure 7. Lithospheric heterogeneity in the region seen in the differences between the subregion-average, phase-velocity curves. The region-average dispersion curves are computed using the optimal-resolution, phase-velocity maps (Section 3). Right: the definition of the entire region (grey) and 4 subregions. The black dots on the map show the knots of the grid used in the phase-velocity tomography and indicate the area sampled by the interstation paths. Left: the phase-velocity anomalies with respect to the region average (top) and the absolute phase velocities in the subregions (bottom). Grey curve: the average of all the measurements in the dataset.

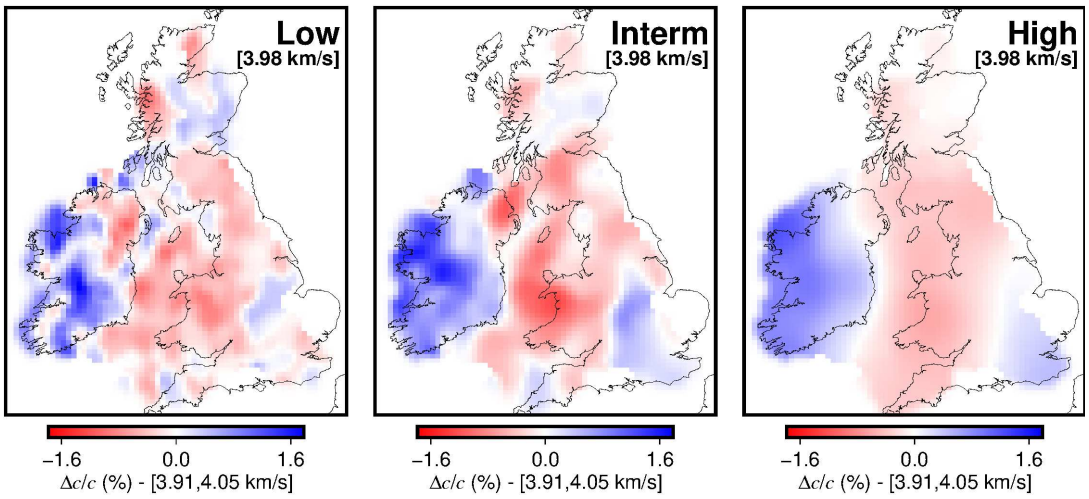

Figure 8. The effect of smoothing on the phase-velocity maps at 44 s period. A map that is not smooth enough (left) fits noise in the data and is dominated by artifacts (noise). An overly smooth map (right) is accurate, at its spatial wavelengths, but at a cost of a decrease in resolution. The intermediately smoothed map is preferable but, at close inspection, does not show an equally optimal regularization everywhere in the region (achieved, instead, in the composite, variable-regularization maps such as in Section 3).

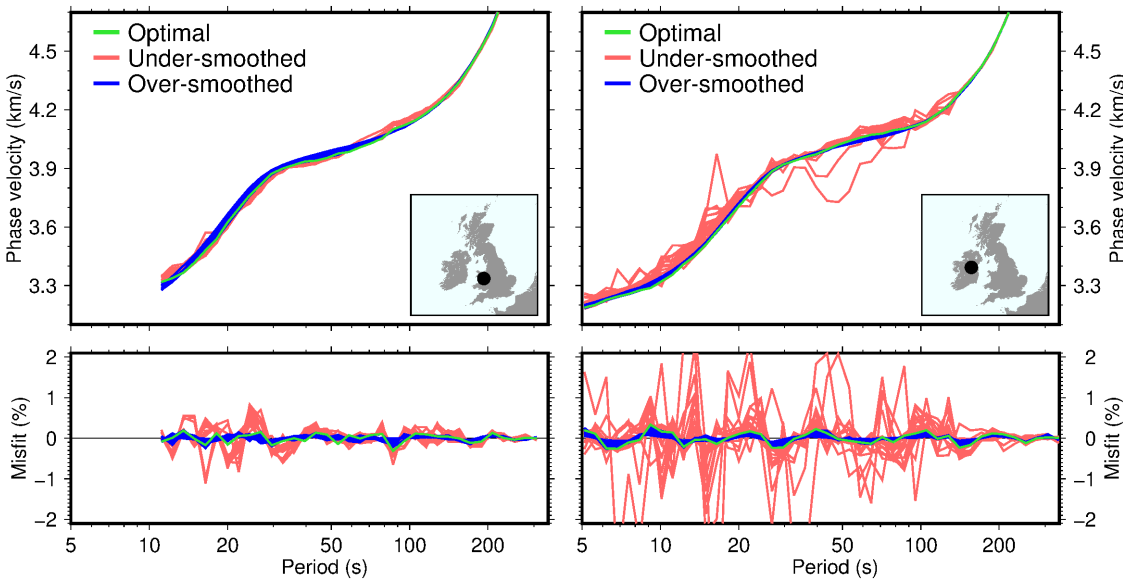

Figure 9. Selection of the optimal inversion smoothing for a point, based on the estimated errors of the phase-velocity curves yielded by phase-velocity maps. Top: phase-velocity curves at two points in Wales and in Ireland. The locations are mapped in the insets. The curves are extracted from phase-velocity maps with different levels of smoothness. Bottom: the period-dependent error of each phase-velocity curve. The error is estimated as the roughness of the phase-velocity curve, which we isolate by fitting it with a synthetic in a very weakly regularized inversion. Red lines: the curves obtained from rough maps. Blue lines: the curves obtained from smooth maps. Green line: the curve chosen as optimal.

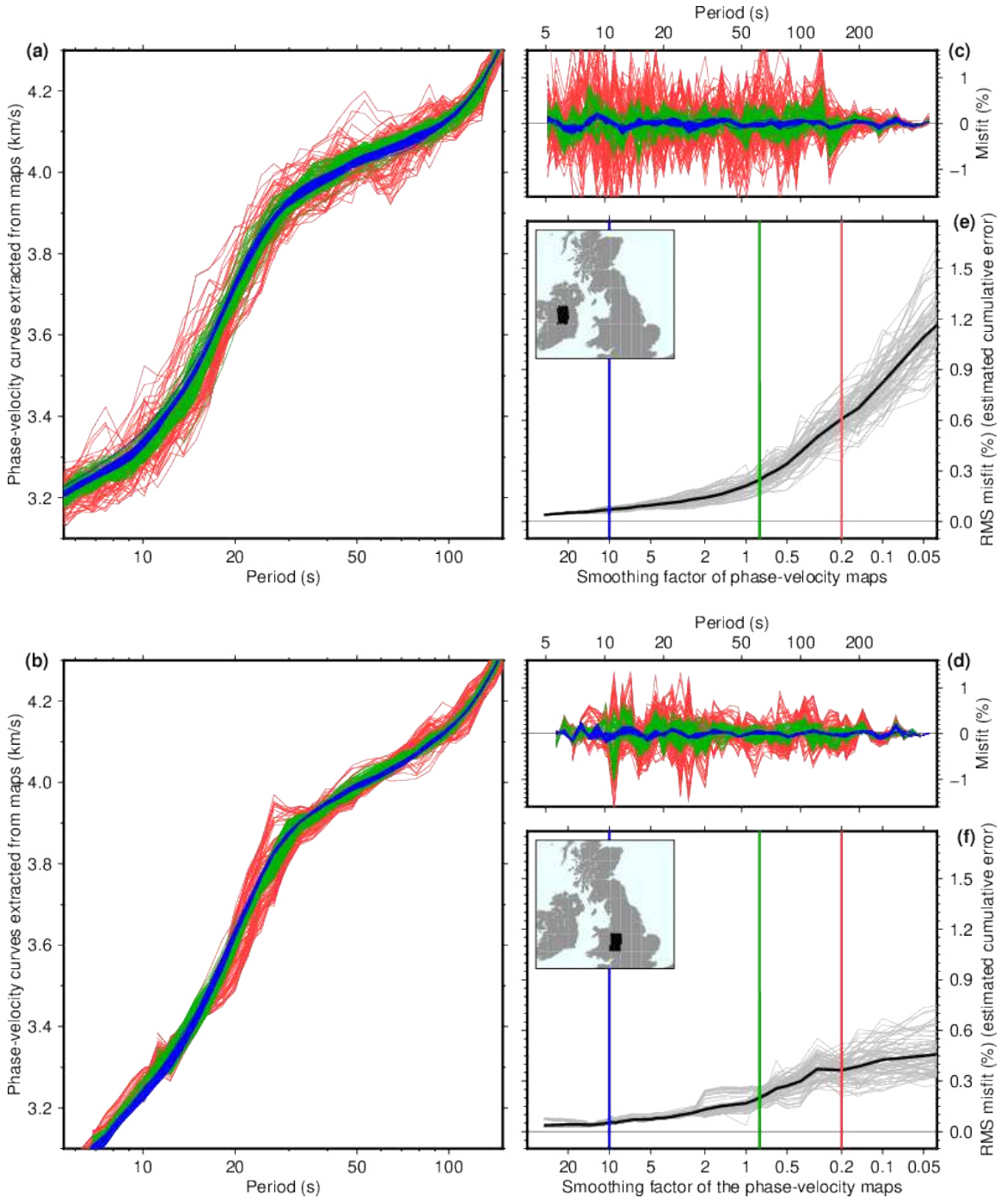

Figure 10. The decrease in the errors of phase-velocity curves with the increasing smoothness of the phase-velocity maps that they are derived from. The errors are estimated from the roughness of the phase-velocity curves. (a), (b): phase-velocity curves extracted from the phase-velocity maps computed with many different smoothing levels, grouped into rough (red), intermediate (green) and smooth (blue), at sets of neighbouring knots in Ireland and Britain (black dots in the maps). (c), (d): the misfits that quantify the period-dependent roughness of the curves and, by inference, their errors. (e), (f): the RMS misfit as a function of the smoothing applied (grey curves). Black: the average across the sets of neighbouring points.

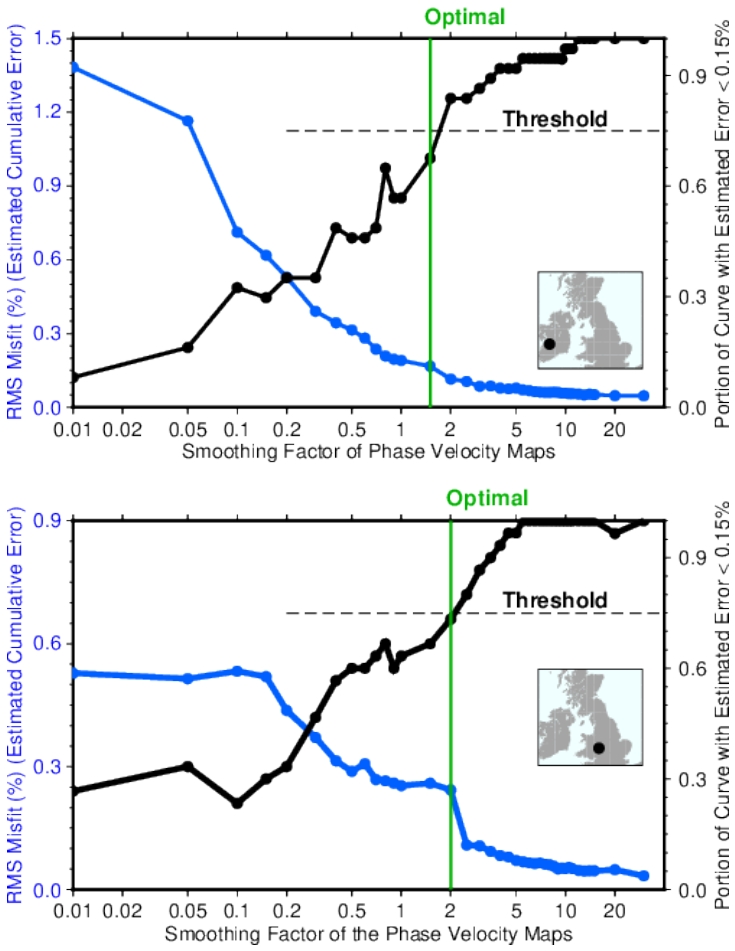

Figure 11. The cumulative error (RMS misfit) and the portion of the phase-velocity curve with the estimated error less than 0.15% (our criterion for selecting the optimal regularization) change nearly monotonically with the smoothing factor of the phase-velocity maps that the curves are extracted from. The two examples are from grid knots in Ireland and Wales.

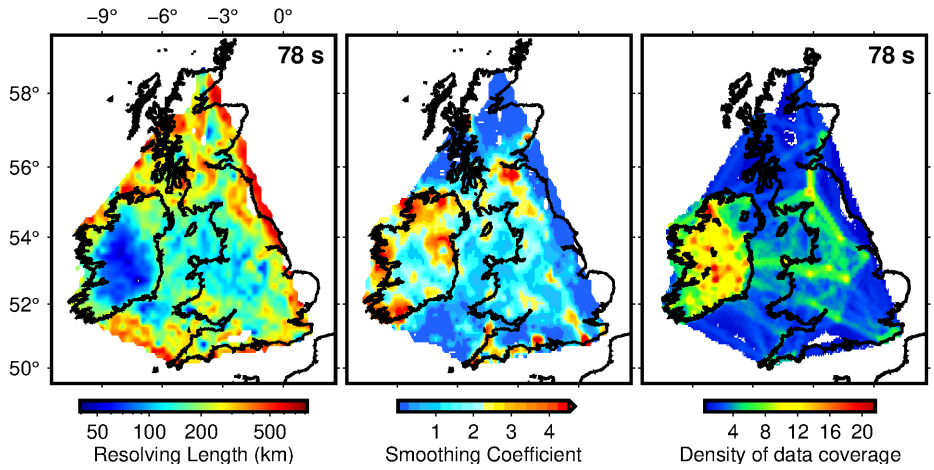

Figure 12. Left: The laterally varying resolving length yielded by the optimal-resolution tomography. A 78-s phase-velocity map is plotted as an example. The resolving length and the parameter value at each point were determined in a series of inversions of the entire system. The optimal width of the averaging kernel—defined as the smallest averaging kernel width such that the error is below a threshold—was found by varying it in a broad range, using varying smoothing parameters. Middle: laterally varying smoothing coefficients determined and applied in the course of the optimal-resolution tomography. Phase-velocity maps at all periods are computed with these smoothing coefficients. Note that the smoothing coefficient value does not scale with the smoothness or the resolving length. Right: the density of the data coverage at 78 s, computed as the sums of the columns of the sensitivity matrix, determined by the path coverage.

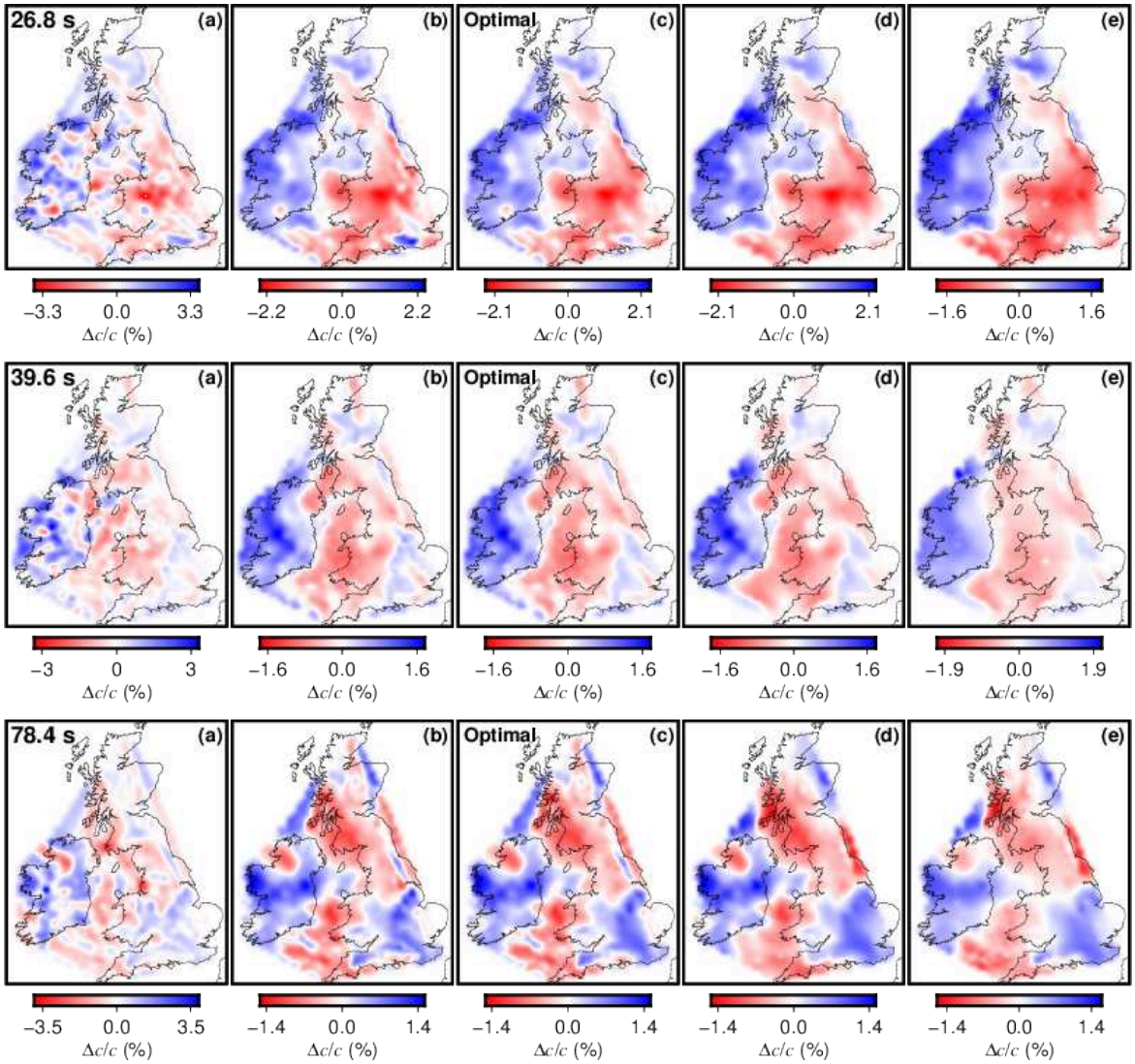

Figure 13. Composite, variable-resolution, phase-velocity maps constructed using different thresholds for acceptable errors of the local dispersion curves. In panel (c) we show the velocity maps with optimal resolution, the maps in (a), (b) and (d), (e) correspond to looser and stricter criteria of selection, respectively, with regard to the error threshold.

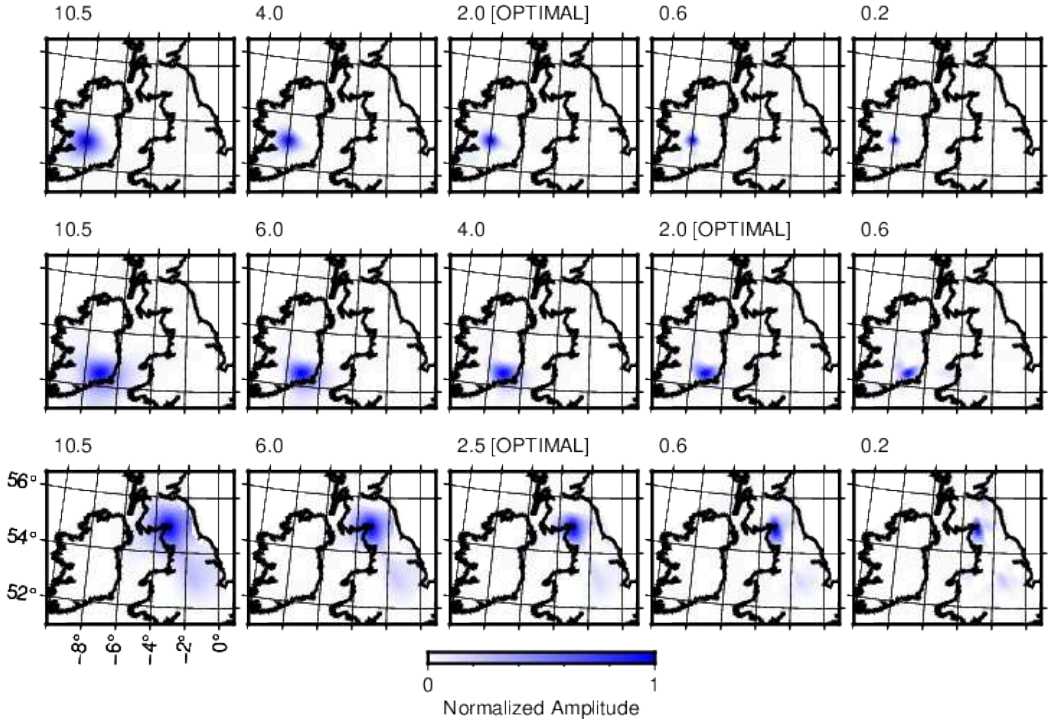

Figure 14. Point-spread functions at three different locations, two in Ireland and one in Scotland, yielded by differently smoothed tomographic test inversions. The smoothing coefficient S is indicated at the top left of each plot.

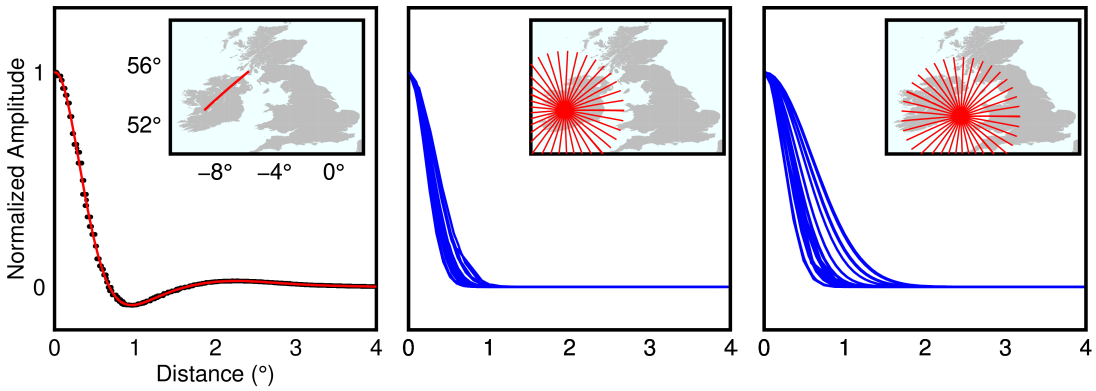

Figure 15. Estimation of the averaging kernel and resolving length. Left: An estimate of the averaging kernel at a point along one azimuth through the point. Anomaly values at grid points along the line shown on the map are matched closely by a bell curve, computed using cubic splines so as to fit the points. Centre, right: illustration of different resolving length for two different locations, given by theoretical gaussian functions computed from the measured half width at half maximum along the full range of densely spaced azimuths. The resolving length is determined as the average of the full width at half maximum given by the curves at the different azimuths.

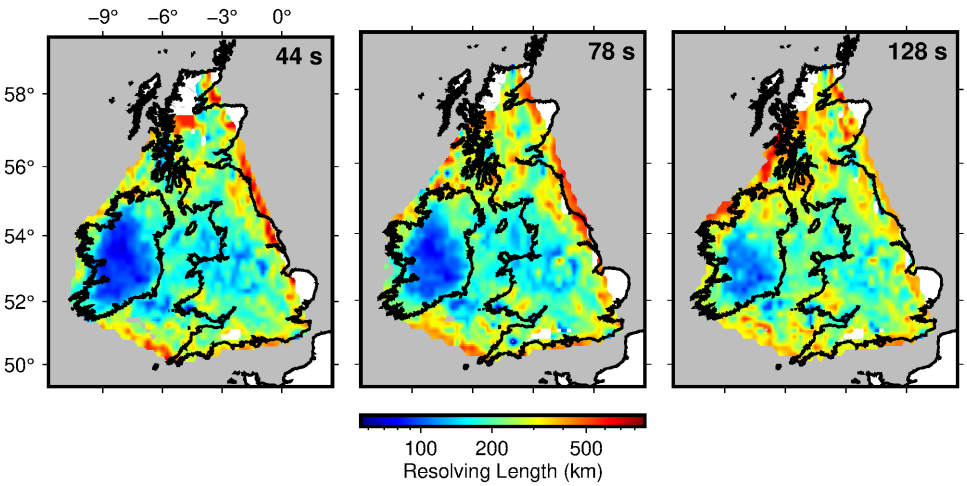

Figure 16. The resolving length of the optimal model is similar at different periods, fulfilling an important condition for the accuracy of the procedure.

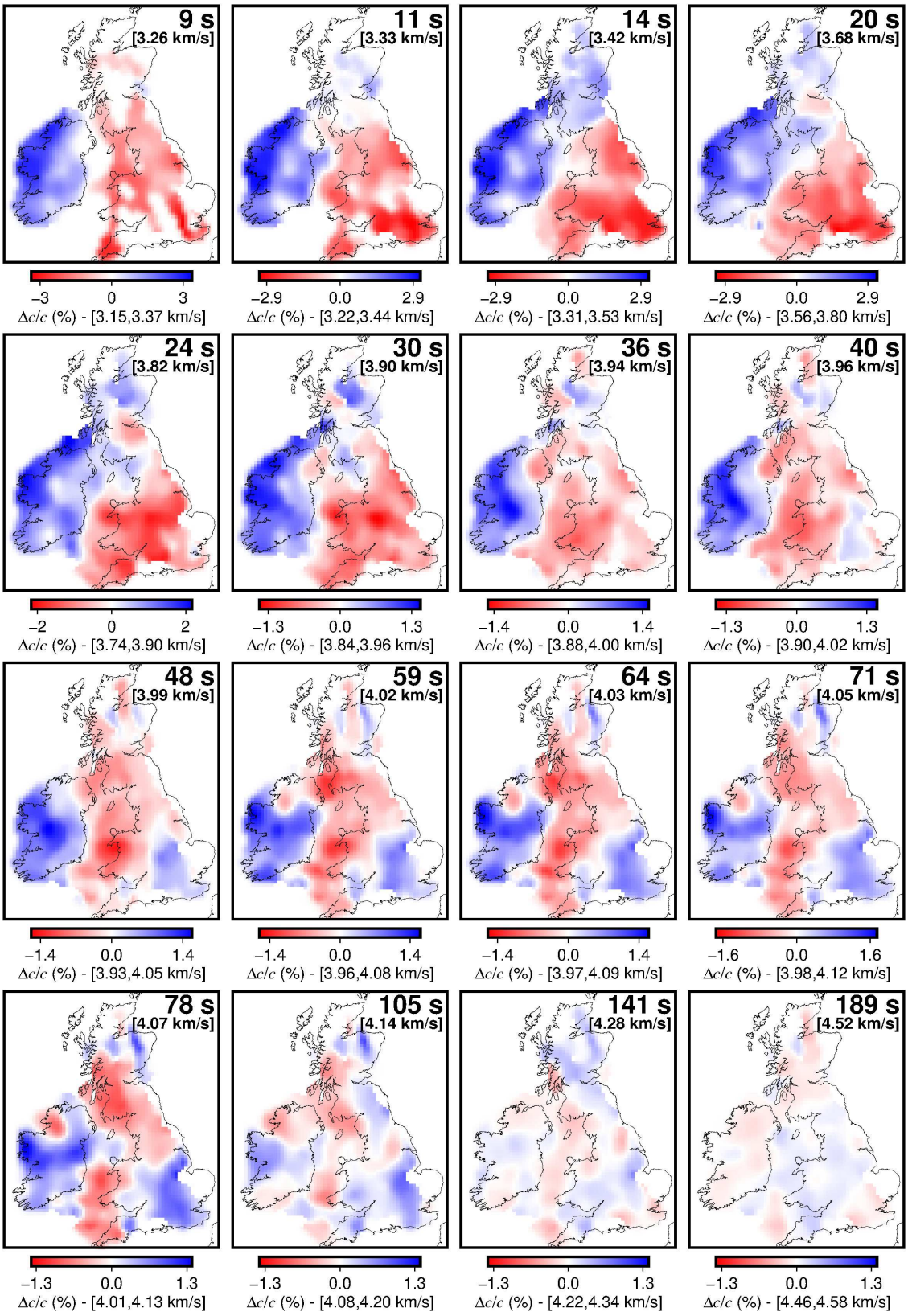

Figure 17. Optimal resolution phase-velocity maps of the region. The phase-velocity anomalies are with respect to the region average, indicated in the top right corner of each frame. The minimum and maximum phase velocity is given below each colour scale. Parts of the maps where the coverage is extremely low (for example, at the edges of the region, where the coverage deteriorates and the data sampling is insufficient to constrain the structure) have been removed from the analysis, based on sensitivity-matrix column sums–accepting only knots with values over a threshold.

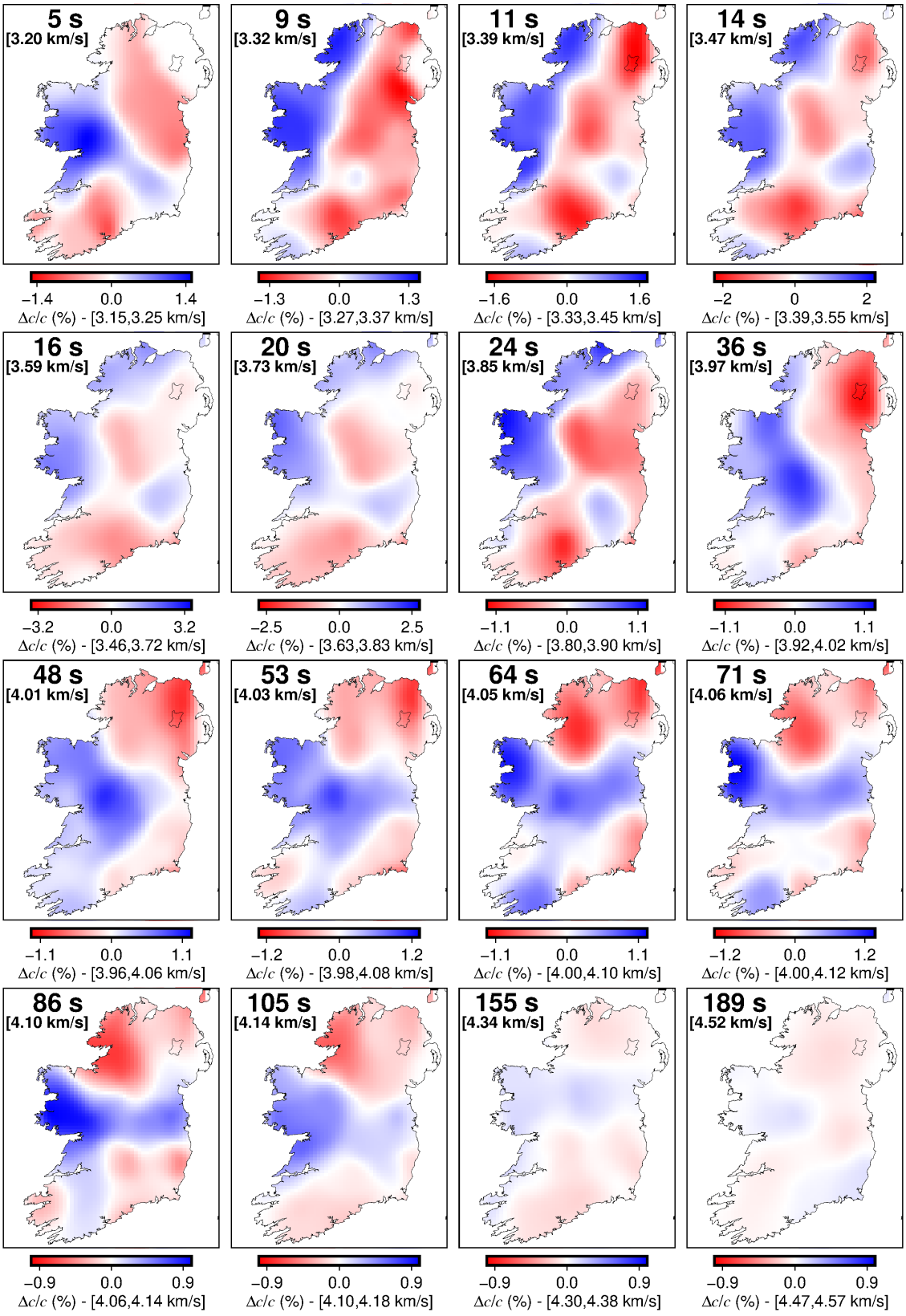

Figure 18. Optimal resolution, phase-velocity maps of Ireland. The phase-velocity anomalies are with respect to the Ireland average, indicated in the top left corner of each frame. The minimum and maximum phase velocity is given below each colour scale.

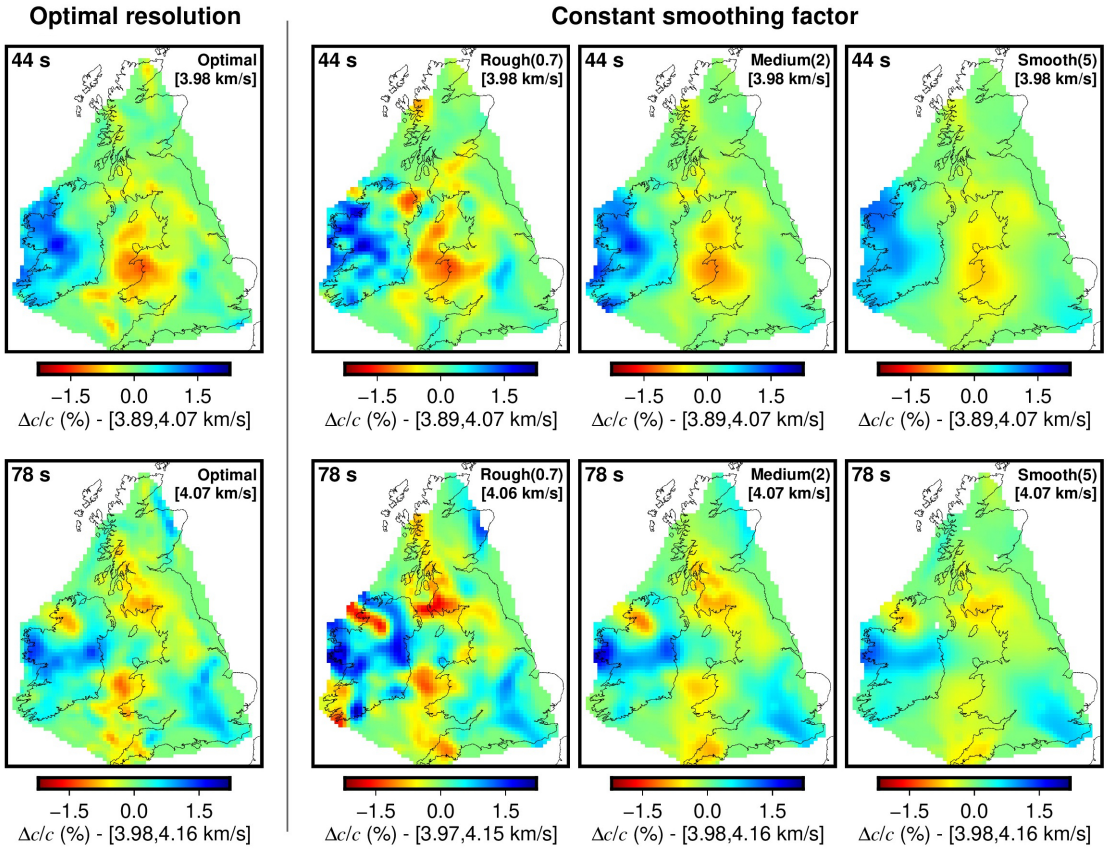

Figure 19. Optimal resolution maps (left column) at two periods (44 s, 78 s) compared with constant-smoothing-factor maps computed using different levels of smoothing (second to fourth columns). The composite, optimal resolution maps display lateral variations in resolution as warranted by the data coverage and errors.

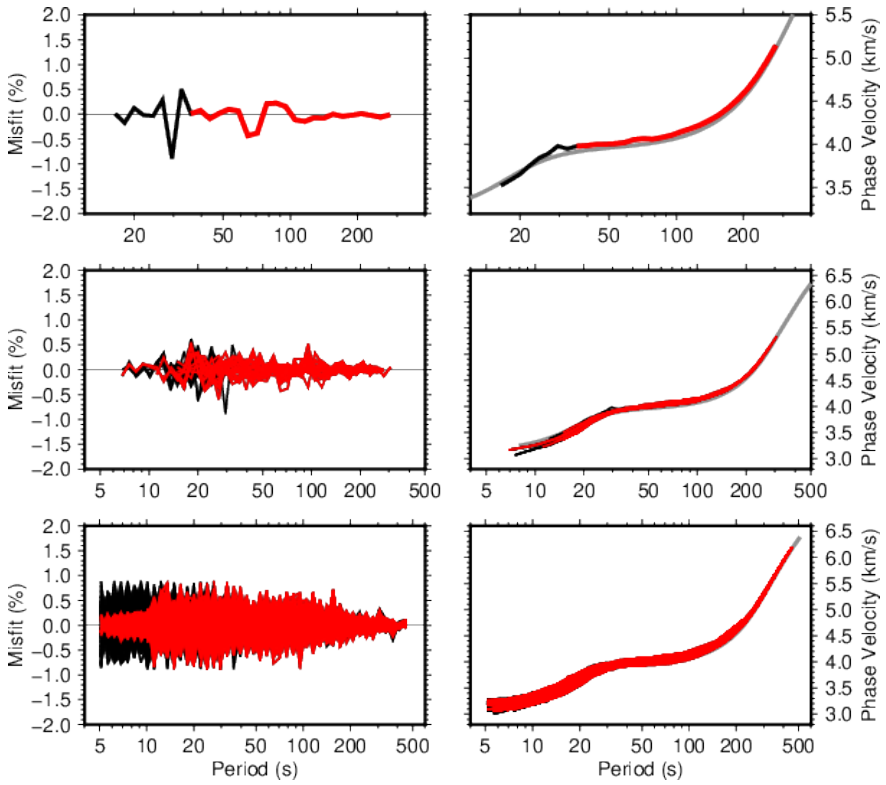

Figure 20. Removal of the "noisy tails" of the phase-velocity curves prior to the point-by-point, 1D inversions for shear-wave velocity structure. The original curve is shown in black, the curve after the removal of the noisy part is shown in red. The phase-velocity curves are shown in the right column, and the period-dependent error estimates (the relative misfits) given by weakly regularised inversions of the curves—in the left column. Top: An example of one phase-velocity curve only. Middle: 30 curves randomly selected from the entire dataset. Bottom: 5000 randomly selected dispersion curves.

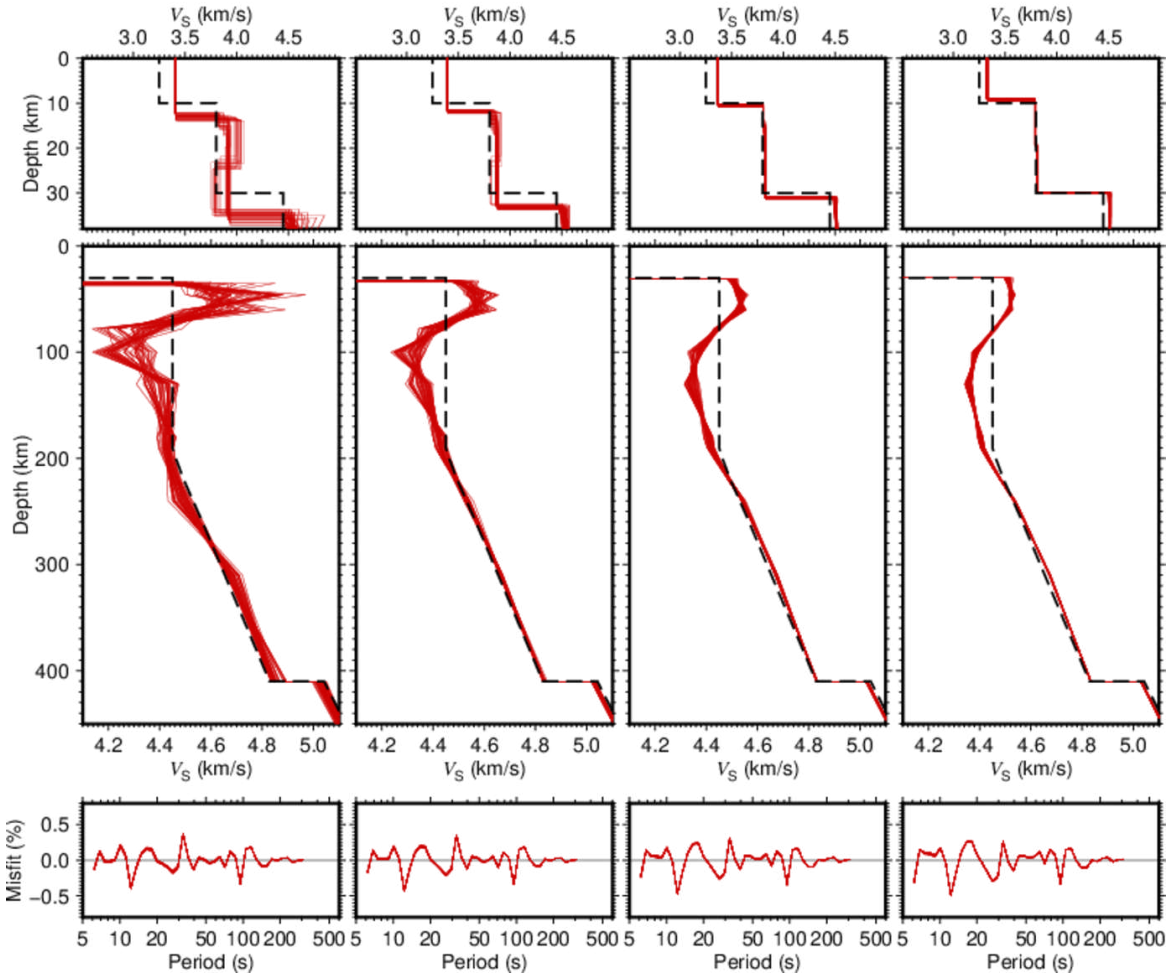

Figure 21. Inversions of the local phase-velocity curve at one grid knot (54.93N, 7.9W) for V_S profiles, with increasing damping from left to right. The reference model is plotted with a dashed black line. Top: a zoom on the crust; middle: the profile from 0 to 450 km; bottom: the phase-velocity misfit. The models yielded by the inversions are non-unique but the robust features in the V_S profile are evident.

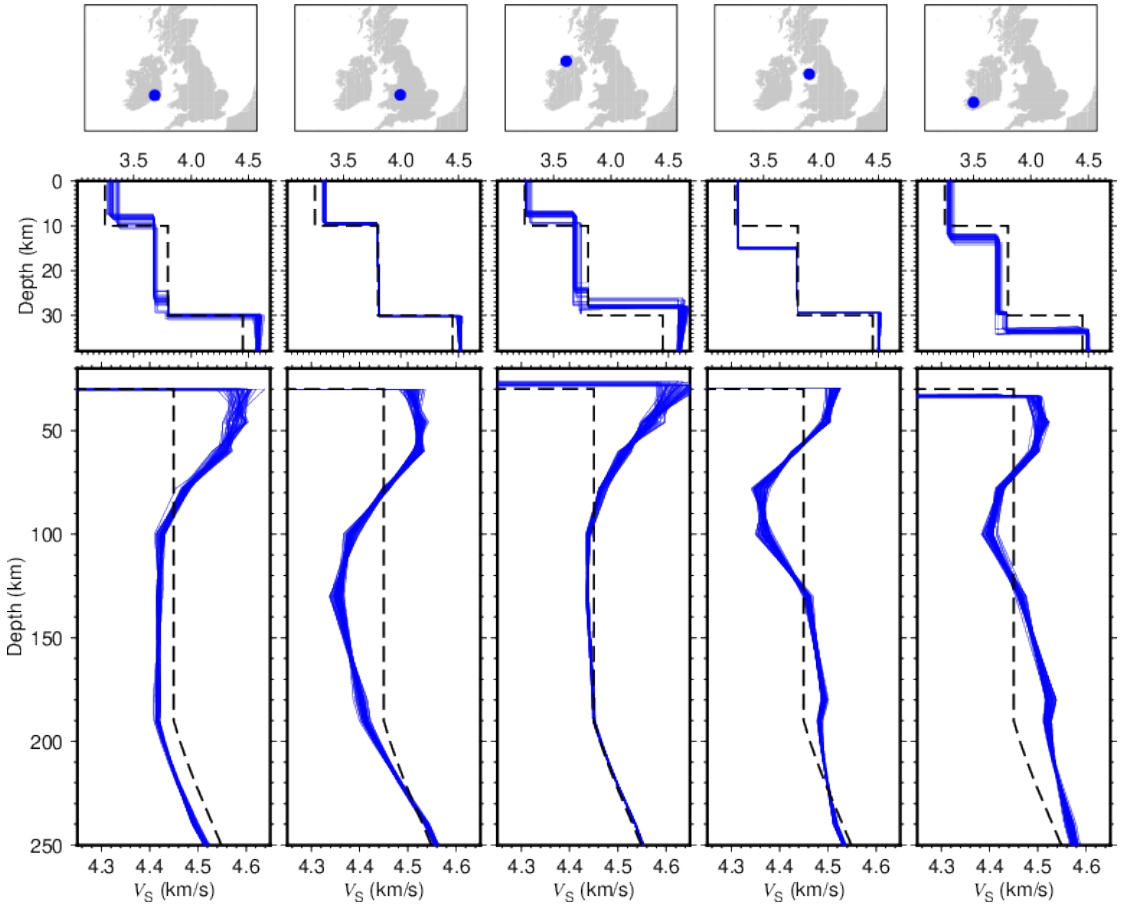

Figure 22. V_S profiles at 5 selected locations in Ireland and Britain.

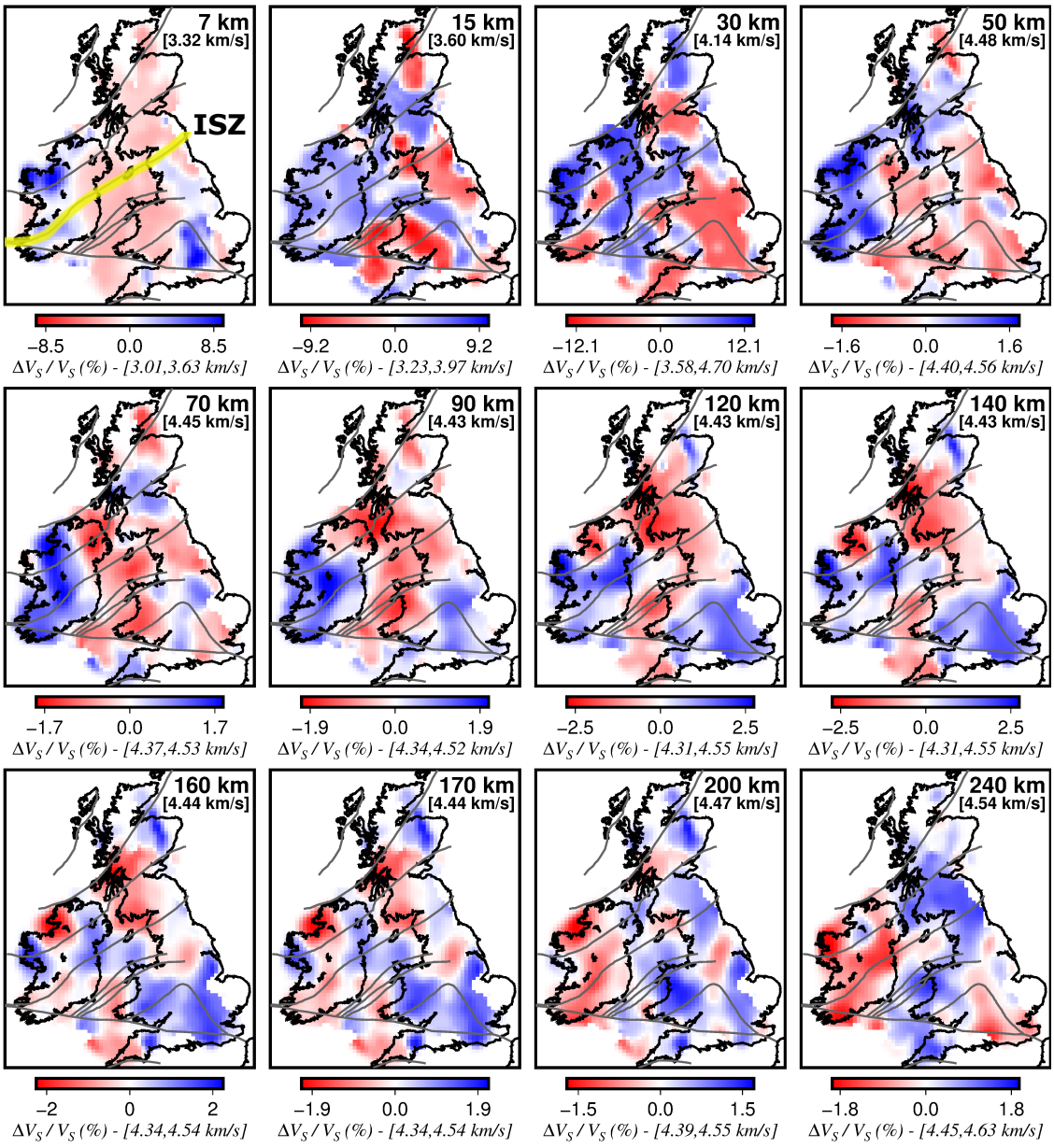

Figure 23. Shear-wave speed anomaly, plotted with respect to the region average at each depth. The grey lines represents major geological boundaries (after Tomlinson et al. 2006). The line highlighted in yellow represents the Iapetus Suture Zone (ISZ) (e.g., Holland & Sanders 2009).

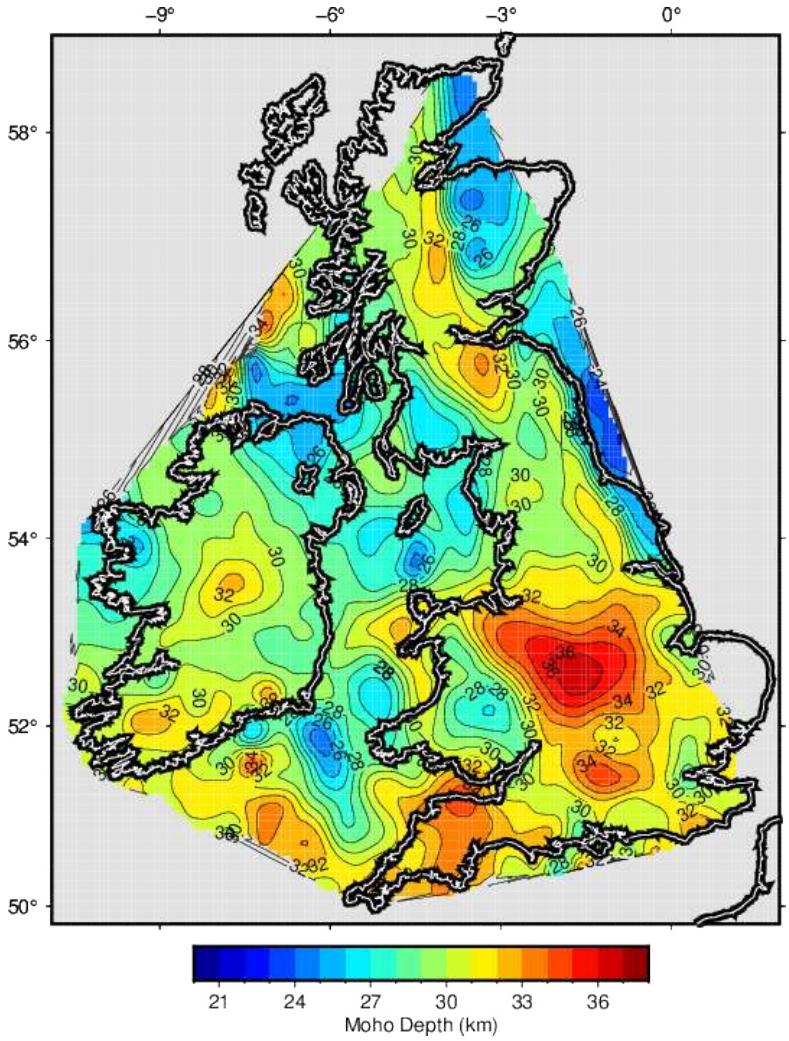

Figure 24. The Moho-depth map of the region yielded by the surface-wave tomography.

APPENDIX A: APPENDIX

A1 Data sources

Until 2010, there was only one broadband (recording both short and long—over 100 s—periods) seismic station on the island of Ireland, the permanent station DSB near Dublin, operated jointly by the GEOFON network (GEOFON Data Centre 1993) and the Dublin Institute for Advanced Studies. Another permanent station, VAL on Valentia Island, was wide-band (periods up to 30 s only) and a few short period stations were operated at different times in the Republic of Ireland (ROI) and Northern Ireland (NI). There were also two temporary deployments of wide-band (30 s) stations: the project ISLE (Irish Seismic Lithospheric Experiment, Landes et al. 2004, 2006; Do et al. 2006; Wawerzinek et al. 2008), with stations in SW Ireland from 2002 to 2005, and the project ISUME (Irish Seismological Upper Mantle Experiment, O'Donnell et al. 2011; Polat et al. 2012) that installed stations across Ireland in 2006, with some of these recording to this day.

In 2010-2012, Ireland Array (Lebedev et al. 2012) deployed 20 broadband (nominally, 120 s, but also recording periods of hundreds of seconds) stations across ROI, with most of the stations recording continuously until present. At the same time, the Irish National Seismic Network (INSN) was established and installed 5 new permanent broadband stations (making it 6 in total, including DSB) (Blake et al. 2012). Broadband stations of the UK Seismograph Network operated by the British Geological Survey (BGS) (Baptie 2018) are distributed across Britain and Northern Ireland and complete the broadband station coverage of the area.

Our dataset includes all the data recorded by the broadband networks in Ireland and all the publicly available data from the broadband stations in Britain. We also used all the data from temporary, wide-band deployments in Ireland, including ISLE, ISUME, WaveOBS (60-s and 30-s stations, Möllhoff & Bean 2016), Dublin Basin array (30-s stations, Licciardi & Piana Agostinetti 2014, 2017) and SIM-CRUST (30-s stations, Piana Agostinetti & Licciardi 2015), and the publicly available data from temporary deployments in Britain, including the Blacknest Array (AWE 2020). In total, our measurements were made on data recorded between 1981 and 2018, but the bulk of the

data is from the last decade. Thanks to the recent growth in the number of stations, especially in Ireland, our dataset provides an unprecedentedly dense data coverage of the entire region (Fig. 3).

A2 Additional figures

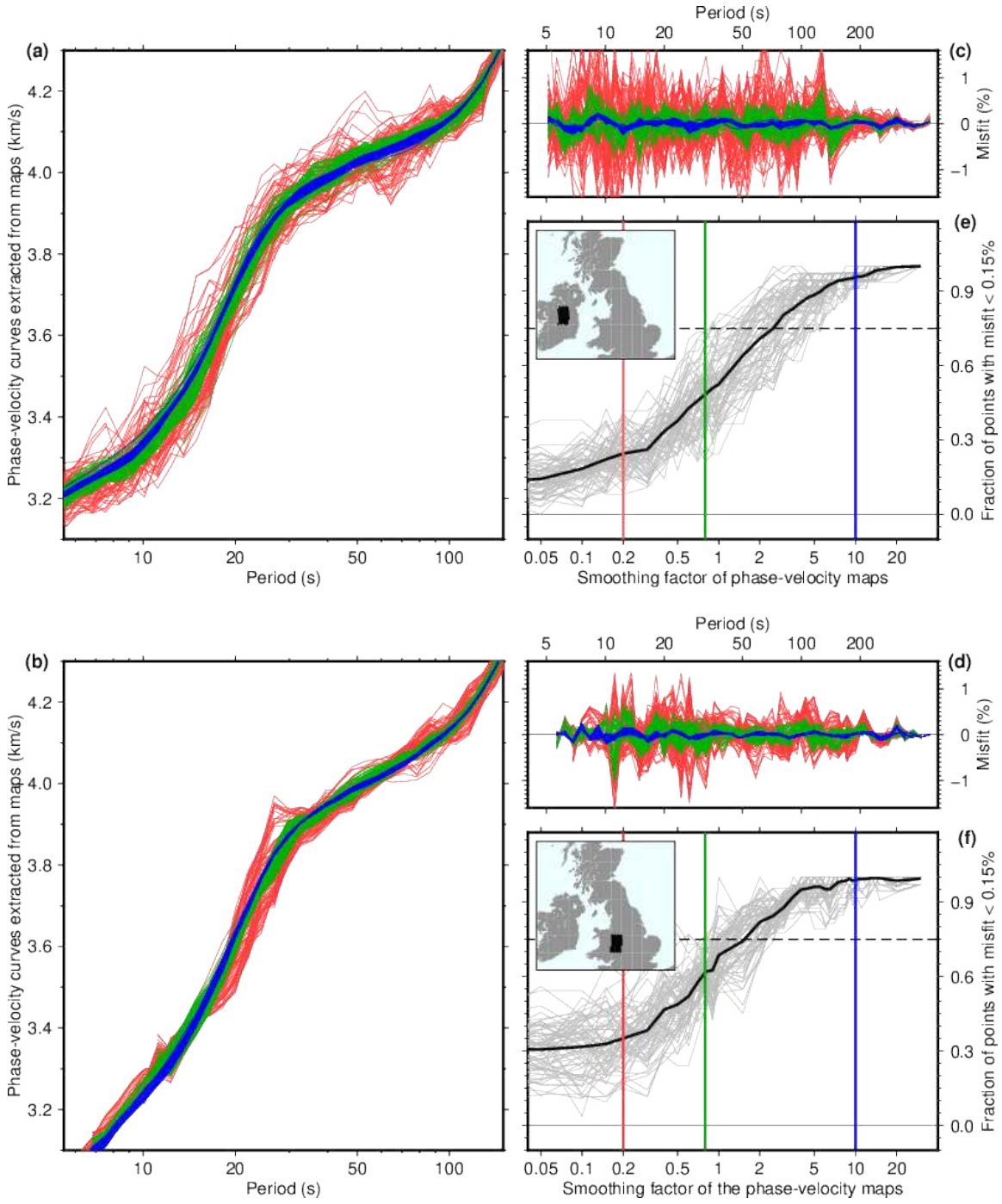

Figure A1. The decrease in the errors of phase-velocity curves with the increasing smoothness of the phase-velocity maps that they are derived from. The errors are estimated from the roughness of the phase-velocity curves (the rapid phase-velocity oscillations with period that could not be explained by any Earth structure). (a), (b): phase-velocity curves extracted from the phase-velocity maps computed with many different smoothing levels, grouped into rough (red), intermediate (green) and smooth (blue), at sets of neighbouring knots in Ireland and Britain (black dots in the maps). (c), (d): the misfits that quantify the period-dependent roughness of the curves and, by inference, their errors. (e), (f): the fraction of points with the misfit lower than 0.15% as a function of the smoothing applied (grey curves). Black: the average across the sets of neighbouring points.

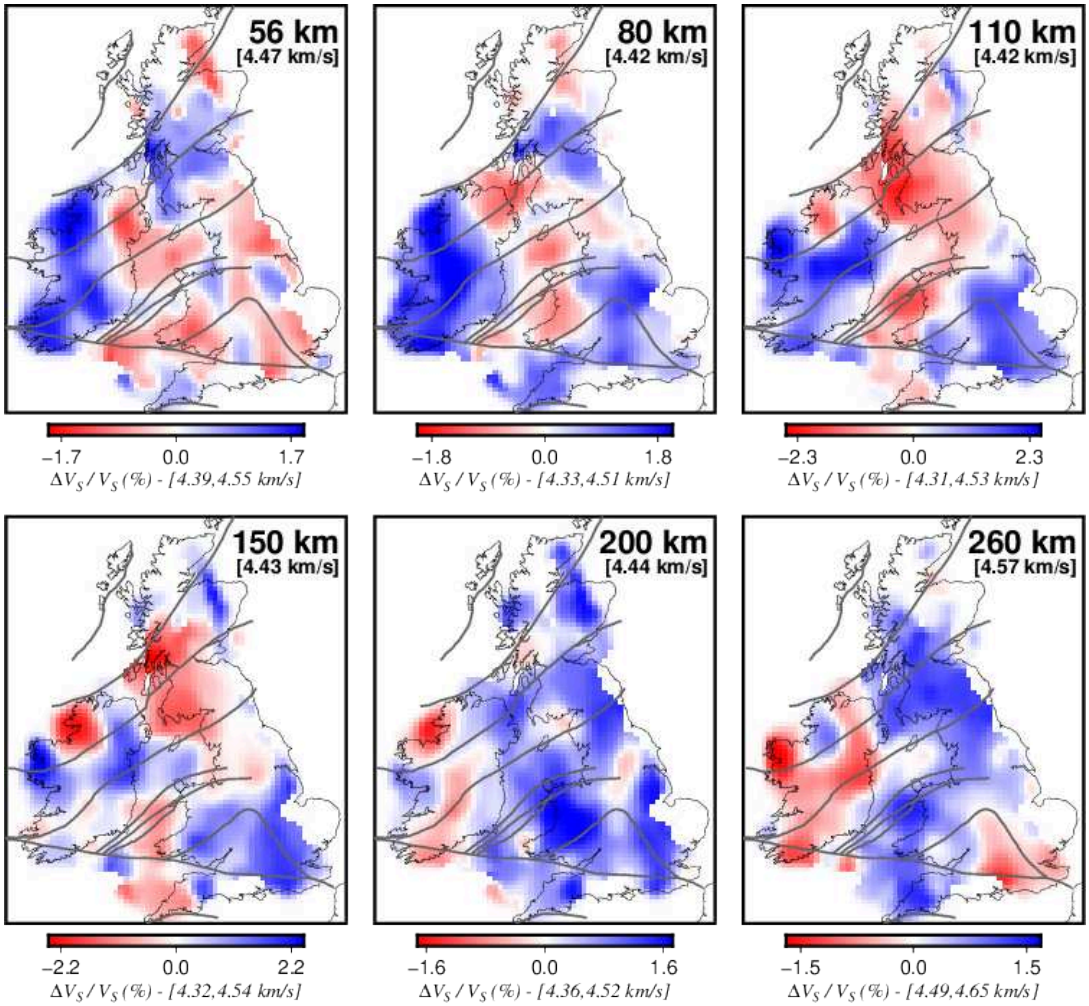

Figure A2. Shear-wave speed anomalies with respect to the global average values at the depths, taken from Schaeffer & Lebedev (2013).

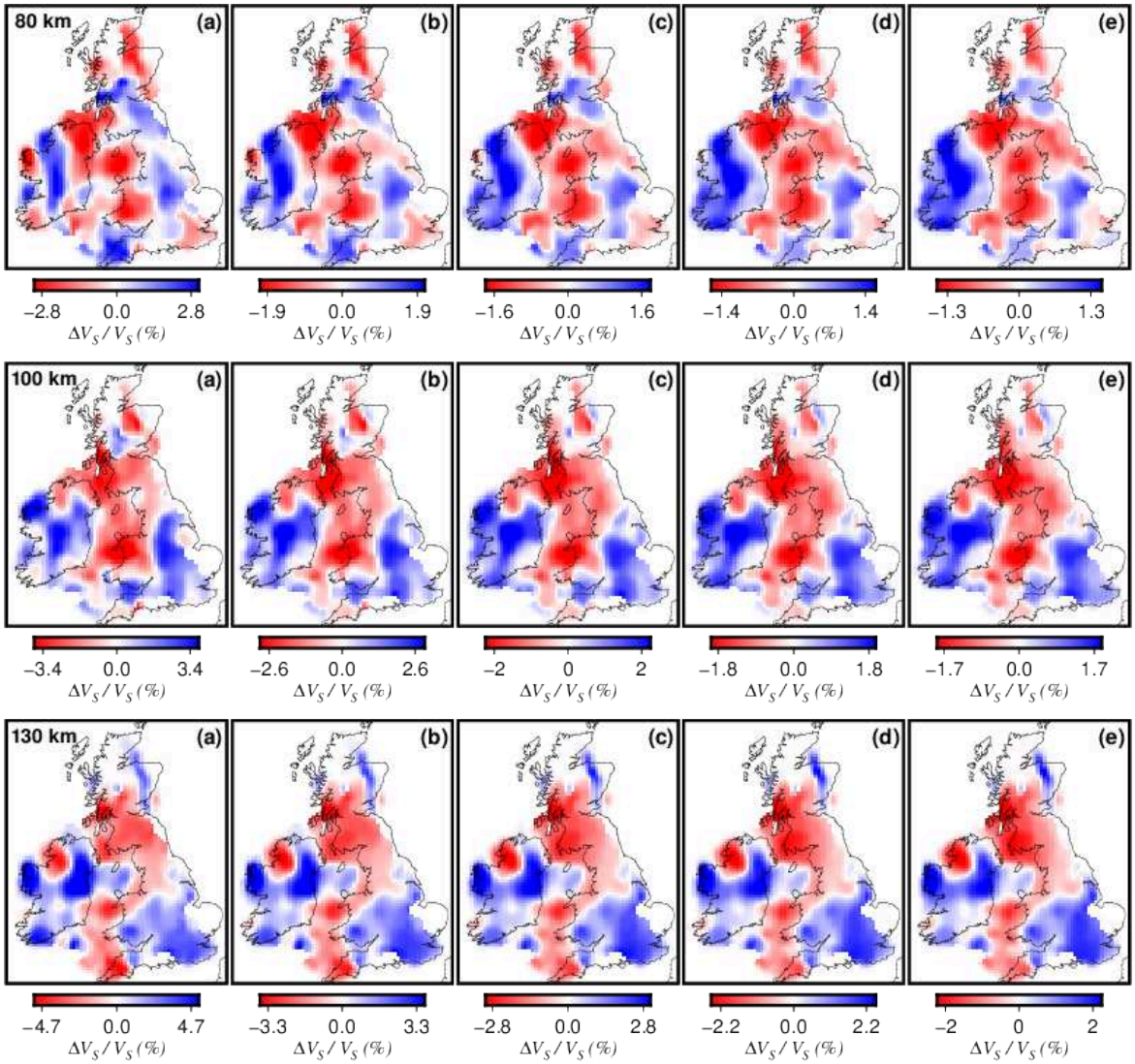

Figure A3. Shear-wave models constructed using 5 different damping levels (increasing from (a) to (e) by a factor of 10) in the point-by-point, 1D inversions, plotted at three different depths.